

FOR OFFICIAL USE ONLY

U.S. DEPARTMENT OF COMMERCE
Patent and Trademark Office

20548
SEARCH REQUEST FORM

Examiner # (Mandatory): 76554 Requester's Full Name: Li, Leo

Art Unit 1645 Location (Bldg/Room#): 8E17 Phone (circle 305-306 308) 8891

Serial Number: 09/107,979 Results Format Preferred (circle): PAPER DISK E-MAIL

Title of Invention _____

Inventors (please provide full names): _____

Earliest Priority Date: _____

Keywords (include any known synonyms registry numbers, explanation of initialisms):

please search

SEQ ID NO = 4

TECH/CHEM. DIVISION
(STIC)

OCT 12 1999

RECEIVED

Search Topic:

Please write detailed statement of the search topic, and the concept of the invention. Describe as specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples of relevant citations, authors, etc., if known. You may include a copy of the abstract and the broadcast or most relevant claim(s).

please Rush it

AC
ANTHONY C. CAPUTA
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 1600

FOR OFFICIAL USE ONLY

STAFF USE ONLY

Searcher: Wang

Searcher Phone #: _____

Searcher Location: _____

Date Picked Up: 10/13

Date Completed: _____

Clerical Prep Time: _____

Terminal Time: 21

Number of Databases: _____

Type of Search

____ N.A. Sequence

1 A.A. Sequence

____ Structure (#)

____ Bibliographic

____ Litigation 1

____ Fulltext

____ Procurement

____ Other

Vendors (include cost where applicable)

____ STN

____ Questel/Orbit

____ Lexis/Nexis

____ WWW/Internet

____ In-house sequence systems (list)

____ Dialog

____ Dr. Link

____ Westlaw

✓ Other (specify) CompuGen

BEST AVAILABLE COPY

THIS PAGE BLANK (USPTO)

Run	Time (min)	Area	Height	Area%	Height%	Ident
44	113.3	41.0	280	1	1	W09369
45	111.5	40.3	422	1	1	R43651

ALIGNMENTS

44	113.5	280	1	W09369	Human neuregulin G
45	111.5	422	1	R43651	Rat NR5F. New reco

AC	W57017;
DC	10-MAY-1999 (first entry)
DE	Mouse neuregulin related ligand NRG3.
DE	Neuregulin related ligand; NRG3; mouse; ErbB4 receptor;
KW	signal transduction; nervous system disorder; neurodegeneration;
KW	neuropathy; therapy; diagnosis.

KW signal transduction; nervous system disorder; neurodegeneration;
KW neuropathy; therapy; diagnosis.

OS	EH	Mus Key	Mus sp.	Education/Qualifiers
OS	EH	Mus Key	Mus sp.	Education/Qualifiers

ET
Domain

...362

/note= "extracellular domain, specifically claim

FT	Region	56. .91	
FT		/note= "hydrophobic region"	
FT	Region	105. .286	
FT		/note= "mucin-like Ser/Thr-rich region, contains sites for O-linked glycosylation"	

Result No.	Query	Score	Query			ID	Description
			Match	Length	DB		
1	277	100.0	713	1	W97617	Mouse neuroregulin r	
2	277	100.0	362	1	W97620	Mouse neuroregulin r	
3	277	100.0	720	1	W97618	Human neuroregulin r	
4	277	100.0	360	1	W97621	Human neuroregulin r	
5	277	100.0	47	1	W97622	Human neuroregulin r	
6	277	100.0	696	1	W97619	Human neuroregulin r	
7	277	100.0	157	1	Y05451	Human neuroregulin-l	
8	277	100.0	720	1	Y05452	Human hereregulin-li	
9	116.5	42.1	52	1	W03182	Neu differentiation	
10	113.5	41.0	263	1	R28537	GGF2BPP2.CDS prote	
11	113.5	41.0	669	1	R29570	Human hereregulin-al	
12	113.5	41.0	63	1	R46918	EGFL2. Glial growth	
13	113.5	41.0	88	1	R46921	EGFL5. Glial growth	
14	113.5	41.0	83	1	R46922	EGFL6. Glial growth	
15	113.5	41.0	63	1	R55659	EGFL2. Glial mitoge	
16	113.5	41.0	88	1	R55662	EGFL5. Glial mitoge	
17	113.5	41.0	83	1	R55663	EGFL6. Glial mitoge	
18	113.5	41.0	263	1	R46896	GGF2BPP2. Glial gr	
19	113.5	41.0	280	1	R46915	GGF2BPP2. Glial gr	
20	113.5	41.0	280	1	R55656	GGF2BPP2. Glial mi	
21	113.5	41.0	263	1	R55689	GGF2BPP2. Glial mi	
22	113.5	41.0	125	1	R68564	Human NDF-alpha3 c	
23	113.5	41.0	462	1	R68562	Human NDF-alpha2b	
24	113.5	41.0	462	1	R68561	Human proNDF-alpha	
25	113.5	41.0	280	1	R67244	Bovine glial cell	
26	113.5	41.0	63	1	R67250	Human epidermal li	
27	113.5	41.0	88	1	R67253	Human epidermal li	
28	113.5	41.0	83	1	R67254	Human epidermal li	
29	113.5	41.0	263	1	R67217	Putative bovine gl	
30	113.5	41.0	63	1	R87461	Epidermal growth f	
31	113.5	41.0	88	1	R87464	Epidermal growth f	
32	113.5	41.0	83	1	R87465	Epidermal growth f	
33	113.5	41.0	280	1	R87455	BPP2 glial growth	
34	113.5	41.0	280	1	R87455	Glial growth facto	
35	113.5	41.0	63	1	R96076	Epidermal growth f	
36	113.5	41.0	83	1	R96080	Epidermal growth f	
37	113.5	41.0	88	1	R96079	Epidermal growth f	
38	113.5	41.0	375	1	W05683	Hereregulin-alpha fu	
39	113.5	41.0	52	1	W03184	Neu differentiation	
40	113.5	41.0	263	1	W03360	Bovine neuroregulin	
41	113.5	41.0	63	1	W03363	EGFL2. Use of neur	
42	113.5	41.0	88	1	W03366	EGFL5. Use of neur	
43	113.5	41.0	83	1	W09367	EGFL6. Use of neur	

Lee

us-09-107-979-4.rag

Thu Oct 14 08:06:37 1999

CC Refsum's disease, abetalipoproteinemia, Tangier disease, Krabbe's
 CC disease, metachromatic leukodystrophy, Fabry's disease and
 CC Dejerine-Sottas syndrome, to treat disease of skeletal muscle of
 CC smooth muscle, such as muscular dystrophy or diseases caused by
 CC skeletal or smooth muscle wasting. The products can also be used
 CC for detection, diagnosis, for the production of transgenic or
 CC knockout animals or for drug screening.
 SQ Sequence 713 AA;

Query Match 100.0%; Score 277; DB 1; Length 713;
 Best Local Similarity 100.0%; Pred. No. 3.6e-21;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HFKPCRDLDLAYCLNDGECFVIETLTGSHKHCRCCKEGYQGVRCDOFL 47
 DB 288 HFKPCRDLDLAYCLNDGECFVIETLTGSHKHCRCCKEGYQGVRCDOFL 334
 |||||||

RESULT 2
 W97620
 ID W97620 standard; Protein: 362 AA.
 AC W97620;
 DT 10-MAY-1999 (first entry)
 DE Mouse neuregulin related ligand NRG3 extracellular domain.
 KW Neuregulin related ligand; NRG3; mouse; ErbB4 receptor;
 KW signal transduction; nervous system disorder; neurodegeneration;
 KW neuropathy; therapy; diagnosis.
 OS Mus sp.
 PN W09902681-A1.
 PD 21-JAN-1999.
 PF 30-JUN-1998; U13411.
 PR 24-JUL-1997; US-899437.
 PR 09-JUL-1997; US-052019.
 PA (GETH) GENENTECH INC.
 PI Godowski PJ, Mark MR, Zhang D;
 WPI: 99-120882/10.
 DR New isolated neuregulin related ligand-3 - used to develop products
 PT for treating nervous system disorders, e.g. stroke, ischaemia,
 PT infection, malignancy, Alzheimer's disease or Down's syndrome
 PS Claim 5(a); Page 62-63; 101pp; English.
 CC This is the extracellular domain (ECD, aa1-362) of murine neuregulin
 CC related ligand NRG3 (see also W97617), a novel member of the
 CC epidermal growth factor (EGF)-like family of protein ligands. NRG3
 CC binds to the ErbB4 receptor, but not to the ErbB2 or ErbB3 receptor,
 CC activates ErbB4 receptor tyrosine phosphorylation. The invention
 CC provides human and murine polypeptides (see also W97618) that have
 CC at least 75% homology to the NRG3 ECD, as well as expression vectors,
 CC host cells and methods for the recombinant production of novel
 CC NRG3s. The NRG3 polypeptides and polynucleotides and can be used to
 CC enhance the survival, proliferation or differentiation of cells
 CC having the ErbB4 receptor in vivo and in vitro. They can be used to
 CC or NRG3-responsive cells, e.g. brain, heart, or kidney cells. In
 CC particular, they can be used to treat diseases which involve neural
 CC cell growth such as demyelination, or damage or loss of glial cells
 CC (e.g. multiple sclerosis). They can be used to treat patients whose
 CC nervous system has been damaged by e.g. trauma, surgery, stroke,
 CC ischaemia, infection, metabolic disease, nutritional deficiency,
 CC malignancy, or toxic agents. NRG3 can also be used to treat
 CC motor neuron disorders such as amyotrophic lateral sclerosis (Lou
 CC Gehrig's disease), Bell's palsy, conditions involving spinal
 CC muscular atrophy or paralysis, neurodegenerative disorders such as
 CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple
 CC sclerosis, Huntington's chorea, Down's syndrome, nerve deafness,
 CC and Meniere's disease. They can also be used to treat neuropathies
 CC associated with systemic disease including post-polio syndrome,
 CC hereditary neuropathies including Charcot-Marie-Tooth disease,
 CC Refsum's disease, abetalipoproteinemia, Tangier disease, Krabbe's
 CC disease, metachromatic leukodystrophy, Fabry's disease and
 CC Dejerine-Sottas syndrome, to treat disease of skeletal muscle of
 CC smooth muscle, such as muscular dystrophy or diseases caused by
 CC skeletal or smooth muscle wasting. The products can also be used

CC for detection, diagnosis, for the production of transgenic or
 CC knockout animals or for drug screening.
 SQ Sequence 362 AA;

Query Match 100.0%; Score 277; DB 1; Length 362;
 Best Local Similarity 100.0%; Pred. No. 1.9e-21;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HFKPCRDLDLAYCLNDGECFVIETLTGSHKHCRCCKEGYQGVRCDOFL 47
 DB 288 HFKPCRDLDLAYCLNDGECFVIETLTGSHKHCRCCKEGYQGVRCDOFL 334
 |||||||

RESULT 3
 W97618
 ID W97618 standard; Protein: 720 AA.
 AC W97618;
 DT 10-MAY-1999 (first entry)
 DE Human neuregulin related ligand NRG3.
 KW Neuregulin related ligand; NRG3; hNRG3B1; human; ErbB4 receptor;
 KW signal transduction; nervous system disorder; neurodegeneration;
 KW neuropathy; therapy; diagnosis.
 OS Homo sapiens.
 PN W09902681-A1.
 PD 21-JAN-1999.
 PF 30-JUN-1998; U13411.
 PR 24-JUL-1997; US-899437.
 PR 09-JUL-1997; US-052019.
 PA (GETH) GENENTECH INC.
 PI Godowski PJ, Mark MR, Zhang D;
 WPI: 99-120882/10.
 DR N-PSDB; X06988.
 DR New isolated neuregulin related ligand-3 - used to develop products
 PT for treating nervous system disorders, e.g. stroke, ischaemia,
 PT infection, malignancy, Alzheimer's disease or Down's syndrome
 PS Claim 5(b); Page 66-69; 101pp; English.
 CC This is the amino acid sequence of human neuregulin related ligand
 CC NRG3, a novel member of the epidermal growth factor (EGF)-like
 CC family of protein ligands that binds to the ErbB4 receptor, but not
 CC to the ErbB2 or ErbB3 receptor, and which activates ErbB4 receptor
 CC tyrosine phosphorylation. The sequence was deduced from the
 CC nucleotide sequence of a cDNA clone (see X06988) from a foetal brain
 CC library. The EGF-like domain of NRG3 is distinct from those of NRG1
 CC or NRG2, and NRG3 displays receptor binding characteristics that are
 CC distinct from those of other neuregulins. An alternatively spliced
 CC form of human NRG3 is provided in W97619. The invention provides
 CC human and murine NRG3 polypeptides (see also W97617), expression
 CC vectors, host cells and methods for the recombinant production of
 CC NRG3s. The NRG3 polypeptides and polynucleotides and can be used to
 CC enhance the survival, proliferation or differentiation of cells
 CC having the ErbB4 receptor in vivo and in vitro. They can be used to
 CC prevent or treat damage to a nerve or damage to other NRG3-expressing
 CC or NRG3-responsive cells, e.g. brain, heart, or kidney cells. In
 CC particular, they can be used to treat diseases which involve neural
 CC cell growth such as demyelination, or damage or loss of glial cells
 CC (e.g. multiple sclerosis). They can be used to treat patients whose
 CC nervous system has been damaged by e.g. trauma, surgery, stroke,
 CC ischaemia, infection, metabolic disease, nutritional deficiency,
 CC malignancy, or toxic agents. NRG3 can also be used to treat

CC motor neuron disorders such as amyotrophic lateral sclerosis (Lou Gehrig's disease), Bell's palsy, conditions involving spinal muscular atrophy or paralysis, neurodegenerative disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, Huntington's chorea, Down's syndrome, nerve deafness, and Meniere's disease. They can also be used to treat neuropathies associated with systemic disease including post-polio syndrome, hereditary neuropathies including Charcot-Marie-Tooth disease, Refsum's disease, abetalipoproteinemia, Tangier disease, Krabbe's disease, metachromatic leukodystrophy, Fabry's disease and Dejerine-Sottas syndrome, to treat disease of skeletal muscle of smooth muscle, such as muscular dystrophy or diseases caused by skeletal or smooth muscle wasting. The products can also be used for detection, diagnosis, for the production of transgenic or knockout animals or for drug screening.

SQ Sequence 720 AA;

Query Match 100.0%; Score 277; DB 1; Length 720;
Best Local Similarity 100.0%; Pred. No. 3.6e-21;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HFPCRDLDLAYCLNDGECFVIETLTGSHKRCCKEGYQGVRCDOFL 47
|||||
Db 286 HFPCRDLDLAYCLNDGECFVIETLTGSHKRCCKEGYQGVRCDOFL 332

RESULT 4

ID W97621
AC W97621 standard; Protein; 360 AA.

DT 10-MAY-1999 (first entry)
DE Human neuregulin related ligand NRG3 extracellular domain.
KW Neuregulin related ligand; NRG3; hNRG3B1; human; ErbB4 receptor;
KW signal transduction; nervous system disorder; neurodegeneration;
KW neuropathy; therapy; diagnosis.
OS Homo sapiens.
OS WO9902681-A1.
PD 21-JAN-1999.
PF 30-JUN-1998; U13411.
PR 24-JUL-1997; US-899437.
PR 09-JUL-1997; US-052019.
PA (GETH) GENENTECH INC.
PI Godowski PJ, Mark MR, Zhang D;
DR WPI: 99-120882/10.

PT New isolated neuregulin related ligand-3 - used to develop products for treating nervous system disorders, e.g. stroke, ischaemia, infection, malignancy, Alzheimer's disease or Down's syndrome
PS Claim 5(a); Page 69-70; 101pp; English.

CC This is the extracellular domain (ECD), aa1-360 of human neuregulin related ligand NRG3 (see also W97618); a novel member of the epidermal growth factor (EGF)-like family of protein ligands. NRG3 binds to the ErbB4 receptor, but not to the ErbB2 or ErbB3 receptor, activates ErbB4 receptor tyrosine phosphorylation. The invention provides human and murine polypeptides (see also W97617) that have at least 75% homology to the NRG3 ECD, as well as expression vectors, host cells and methods for the recombinant production of novel NRG3s. The NRG3 polypeptides and polynucleotides can be used to enhance the survival, proliferation or differentiation of cells having the ErbB4 receptor in vivo and in vitro. They can be used to prevent or treat damage to a nerve or damage to other NRG3-expressing or NRG3-responsive cells, e.g. brain, heart, or kidney cells. In particular, they can be used to treat diseases which involve neural cell growth such as demyelination, or damage or loss of glial cells (e.g. multiple sclerosis). They can be used to treat patients whose nervous system has been damaged by e.g. trauma, surgery, stroke, ischaemia, infection, metabolic disease, nutritional deficiency, malignancy, or toxic agents. NRG3 can also be used to treat motor neuron disorders such as amyotrophic lateral sclerosis (Lou Gehrig's disease), Bell's palsy, conditions involving spinal muscular atrophy or paralysis, neurodegenerative disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, Huntington's chorea, Down's syndrome, nerve deafness, and Meniere's disease. They can also be used to treat neuropathies associated with systemic disease including post-polio syndrome, hereditary neuropathies including Charcot-Marie-Tooth disease, Refsum's disease, abetalipoproteinemia, Tangier disease, Krabbe's disease, metachromatic leukodystrophy, Fabry's disease and Dejerine-Sottas syndrome, to treat disease of skeletal muscle of smooth muscle, such as muscular dystrophy or diseases caused by skeletal or smooth muscle wasting. The products can also be used for detection, diagnosis, for the production of transgenic or knockout animals or for drug screening.

SQ Sequence 360 AA;

CC and Meniere's disease. They can also be used to treat neuropathies associated with systemic disease including post-polio syndrome, hereditary neuropathies including Charcot-Marie-Tooth disease, Refsum's disease, abetalipoproteinemia, Tangier disease, Krabbe's disease, metachromatic leukodystrophy, Fabry's disease and Dejerine-Sottas syndrome, to treat disease of skeletal muscle of smooth muscle, such as muscular dystrophy or diseases caused by skeletal or smooth muscle wasting. The products can also be used for detection, diagnosis, for the production of transgenic or knockout animals or for drug screening.

SQ Sequence 360 AA;

Query Match 100.0%; Score 277; DB 1; Length 360;
Best Local Similarity 100.0%; Pred. No. 1.9e-21;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HFPCRDLDLAYCLNDGECFVIETLTGSHKRCCKEGYQGVRCDOFL 47
|||||
Db 286 HFPCRDLDLAYCLNDGECFVIETLTGSHKRCCKEGYQGVRCDOFL 332

RESULT 5

ID W97622
AC W97622 standard; Protein; 47 AA.

DT 10-MAY-1999 (first entry)
DE Human neuregulin related ligand NRG3 EGF-like domain.
KW Neuregulin related ligand; NRG3; hNRG3B1; human; ErbB4 receptor;
KW signal transduction; nervous system disorder; neurodegeneration;
KW neuropathy; therapy; diagnosis; epidermal growth factor; EGF;
KW immunoadhesin.
OS Homo sapiens.
OS WO9902681-A1.
PD 21-JAN-1999.
PF 30-JUN-1998; U13411.
PR 24-JUL-1997; US-899437.
PR 09-JUL-1997; US-052019.
PA (GETH) GENENTECH INC.
PI Godowski PJ, Mark MR, Zhang D;
DR WPI: 99-120882/10.

PT New isolated neuregulin related ligand-3 - used to develop products for treating nervous system disorders, e.g. stroke, ischaemia, infection, malignancy, Alzheimer's disease or Down's syndrome
PS Claim 30; Page 64; 101pp; English.

CC This is the epidermal growth factor (EGF)-like domain of human neuregulin related ligand NRG3 (see also W97618), a novel member of the EGF-like family of protein ligands that binds to the ErbB4 receptor and activates ErbB4 receptor tyrosine phosphorylation. The EGF-1 like domain of NRG3 is distinct from the EGF-like domains of NRG1 and NRG2. The invention provides human and murine polypeptides (see also W97617) that have at least 75% homology to the NRG3 EGF-like domain, as well as expression vectors, host cells and methods for the recombinant production of novel NRG3s. The NRG3 polypeptides and polynucleotides can be used to enhance the survival, proliferation or differentiation of cells having the ErbB4 receptor in vivo and in vitro. They can be used to prevent or treat damage to a nerve or damage to other NRG3-expressing or NRG3-responsive cells, e.g. brain, heart, or kidney cells. In particular, they can be used to treat diseases which involve neural cell growth such as demyelination, or damage or loss of glial cells (e.g. multiple sclerosis). They can be used to treat patients whose nervous system has been damaged by e.g. trauma, surgery, stroke, ischaemia, infection, metabolic disease, nutritional deficiency, malignancy, or toxic agents. NRG3 can also be used to treat motor neuron disorders such as amyotrophic lateral sclerosis (Lou Gehrig's disease), Bell's palsy, conditions involving spinal muscular atrophy or paralysis, neurodegenerative disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, Huntington's chorea, Down's syndrome, nerve deafness, and Meniere's disease. They can also be used to treat neuropathies associated with systemic disease including post-polio syndrome, hereditary neuropathies including Charcot-Marie-Tooth disease, Refsum's disease, abetalipoproteinemia, Tangier disease, Krabbe's disease, metachromatic leukodystrophy, Fabry's disease and Dejerine-Sottas syndrome, to treat disease of skeletal muscle of smooth muscle, such as muscular dystrophy or diseases caused by skeletal or smooth muscle wasting. The products can also be used for detection, diagnosis, for the production of transgenic or knockout animals or for drug screening.

SQ Sequence 360 AA;

CC Refsum's disease, abetalipoproteinemia, Tangier disease, Krabbe's
 CC disease, metachromatic leukodystrophy, Fabry's disease and
 CC Dejerine-Sottas syndrome, to treat disease of skeletal muscle of
 CC smooth muscle, such as muscular dystrophy or diseases caused by
 CC skeletal or smooth muscle wasting. The products can also be used
 CC for detection, diagnosis, for the production of transgenic or
 CC knockout animals or for drug screening. A claimed immunoadhesin
 CC comprises the human NRG3 EGF-like domain fused to an immunoglobulin
 CC sequence. 47 AA;
 SQ

Query Match 100.0%; Score 277; DB 1; Length 47;
 Best Local Similarity 100.0%; Pred. No. 3e-22;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HFPCRDLDLAYCLNDGECFVIETLTGSHKRCCKEGYGVRCDOFL 47
 |||||
 DB 1 HFPCRDLDLAYCLNDGECFVIETLTGSHKRCCKEGYGVRCDOFL 47

RESULT 6
 W97619
 ID W97619 standard; Protein; 696 AA.
 AC W97619;
 DT 10-MAY-1999 (first entry)
 DE Human neuregulin related ligand NRG3 (splice variant).
 KW Neuregulin related ligand; NRG3; hNRG3B1; human; ErbB4 receptor;
 KW signal transduction; nervous system disorder; neurodegeneration;
 KW neuropathy; therapy; diagnosis; splice variant.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Domain 1..360
 FT /note= "extracellular domain, specifically claimed
 FT in Claim 5(a)"
 FT Region 66..91
 FT /note= "hydrophobic region"
 FT Region 101..284
 FT /note= "mucin-like Ser/Thr-rich region, contains
 FT sites for O-linked glycosylation"
 FT Domain 285..354
 FT /note= "EGF-like domain"
 FT Domain 356..394
 FT /note= "transmembrane domain"
 FT W09902681-AL.
 PN 21-JAN-1999.
 PD 30-JUN-1998; U13411.
 PF 24-JUL-1997; US-899437.
 PR 09-JUL-1997; US-052019.
 PA (GETH) GENENTECH INC.
 PI Godowski P, Mark MR, Zhang D;
 DR WPI: 99-12082/10.
 DR N-PSDB: X06989.
 PT New isolated neuregulin related ligand-3 - used to develop products
 PT for treating nervous system disorders, e.g. stroke, ischaemia,
 PT infection, malignancy, Alzheimer's disease or Down's syndrome
 PS Example 1: Page 78-81; 101pp; English.
 CC This is the amino acid sequence of splice variant hNRG3B2 of human
 CC neuregulin related ligand NRG3, a novel member of the epidermal
 CC growth factor (EGF)-like family of protein ligands that binds to
 CC the ErbB4 receptor, but not to the ErbB2 or ErbB3 receptor, and
 CC which activates ErbB4 receptor tyrosine phosphorylation. The
 CC sequence was deduced from the nucleotide sequence of a cDNA clone
 CC (see X06989) from a foetal brain library. hNRG3B2 lacks amino
 CC acids 329-552 of hNRG3B1 (see W97618) but retains the EGF-like
 CC domain and is expected to exhibit biological activity. The invention
 CC provides human and murine NRG3 polypeptides (see W97617), expression
 CC vectors, host cells and methods for the recombinant production of
 CC NRG3s. The NRG3 polypeptides and polynucleotides and can be used to
 CC enhance the survival, proliferation or differentiation of cells
 CC having the ErbB4 receptor in vivo and in vitro. They can be used to
 CC prevent or treat damage to a nerve or damage to other NRG3-expressing
 CC or NRG3-responsive cells, e.g. brain, heart, or kidney cells. In

CC particular, they can be used to treat diseases which involve neural
 CC cell growth such as demyelination, or damage or loss of glial cells
 CC (e.g. multiple sclerosis). They can be used to treat patients whose
 CC nervous system has been damaged by e.g. trauma, surgery, stroke,
 CC ischaemia, infection, metabolic disease, nutritional deficiency,
 CC malignancy, or toxic agents. NRG3 can also be used to treat
 CC motor neuron disorders such as amyotrophic lateral sclerosis (Lou
 CC Gehrig's disease). Bell's palsy, conditions involving spinal
 CC muscular atrophy or paralysis, neurodegenerative disorders such as
 CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple
 CC sclerosis, Huntington's chorea, Down's syndrome, nerve deafness,
 CC and Meniere's disease. They can also be used to treat neuropathies
 CC associated with systemic disease including post-polio syndrome,
 CC hereditary neuropathies including Charcot-Marie-Tooth disease,
 CC Refsum's disease, abetalipoproteinemia, Tangier disease, Krabbe's
 CC disease, metachromatic leukodystrophy, Fabry's disease and
 CC Dejerine-Sottas syndrome, to treat disease of skeletal muscle of
 CC smooth muscle, such as muscular dystrophy or diseases caused by
 CC skeletal or smooth muscle wasting. The products can also be used
 CC for detection, diagnosis, for the production of transgenic or
 CC knockout animals or for drug screening.
 SQ Sequence 696 AA;

Query Match 100.0%; Score 277; DB 1; Length 696;
 Best Local Similarity 100.0%; Pred. No. 3.5e-21;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HFPCRDLDLAYCLNDGECFVIETLTGSHKRCCKEGYGVRCDOFL 47
 |||||
 DB 286 HFPCRDLDLAYCLNDGECFVIETLTGSHKRCCKEGYGVRCDOFL 332

RESULT 7
 Y05451
 ID Y05451 standard; Protein; 157 AA.
 AC Y05451;
 DT 06-JUL-1999 (first entry)
 DE Human heregulin-like factor sequence.
 KW Human heregulin-like factor; HLF; cell growth regulator; diagnosis;
 KW neural system disorder; cancer.
 OS Homo sapiens.
 PN W09857989-AL.
 PD 23-DEC-1998.
 PF 16-JUN-1998; U12403.
 PR 17-JUN-1997; US-049942.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (GEOU) UNIV GEORGETOWN.
 PI Hijazi MM, King CR, Ruben SM, Young P;
 DR WPI: 99-095327/06.
 DR N-PSDB: X36423.
 PT New isolated heregulin-like factor - used to develop products for
 PT the diagnosis and treatment of disorders involving regulation of
 PT cell growth, particularly cancers
 PS Claim 17; Page 86-87; 118pp; English.
 CC This sequence is the human heregulin-like factor (HLF) of the
 CC invention. The HLF is involved in the regulation of cell growth.
 CC Detection of different levels of expression of the HLF gene can be used
 CC for the diagnosis of disorders, e.g. in the neural system. In
 CC particular, detection of different levels of HLF gene expression in cells
 CC or body fluid of an individual can be used for diagnosing cancer. The
 CC products can also be used in the treatment of disorders involving
 CC abnormal levels of HLF activity.
 SQ Sequence 157 AA;

Query Match 100.0%; Score 277; DB 1; Length 157;
 Best Local Similarity 100.0%; Pred. No. 9e-22;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HFPCRDLDLAYCLNDGECFVIETLTGSHKRCCKEGYGVRCDOFL 47
 |||||
 DB 31 HFPCRDLDLAYCLNDGECFVIETLTGSHKRCCKEGYGVRCDOFL 77

04-JUN-1997 (first entry)	04-JUN-1997 (first entry)
NEW differentiation factor/hereregulin-alpha/beta form EGF-like domain.	NEW differentiation factor/hereregulin-alpha/beta form EGF-like domain.
NDF; new differentiation factor; heregulin; epidermal growth factor;	NDF; new differentiation factor; heregulin; epidermal growth factor;
EGF; colon epithelial cell proliferation; Schwann cell; nerve;	EGF; colon epithelial cell proliferation; Schwann cell; nerve;
damage; colitis; ulcer.	damage; colitis; ulcer.
Synthetic.	Synthetic.
W09631599-Al.	W09631599-Al.
PD 10-OCT-1996.	PD 10-OCT-1996.
27-MAR-1996; U04262.	27-MAR-1996; U04262.
06-APR-1995; US-417640.	06-APR-1995; US-417640.
(AMGE-) AMGEN INC.	(AMGE-) AMGEN INC.
Cranchan JF, Hara S, Lu HS, Mayer JP, Yoshinaga SK;	Cranchan JF, Hara S, Lu HS, Mayer JP, Yoshinaga SK;
WPI: 96-465022/46.	WPI: 96-465022/46.
Peptides-) derived from neu differentiation factor/hereregulin	Peptides-) derived from neu differentiation factor/hereregulin
proteins - specifically from epidermal growth factor-like domain,	proteins - specifically from epidermal growth factor-like domain,
stimulate proliferation of colon epithelial cells and Schwann cells	stimulate proliferation of colon epithelial cells and Schwann cells
Claim 1; Page 27; 37pp; English.	Claim 1; Page 27; 37pp; English.
The peptides W05182-W05185 are based on neu differentiation factor	The peptides W05182-W05185 are based on neu differentiation factor
(NDF)/hereregulin alpha and beta form EGF-like domains in various	(NDF)/hereregulin alpha and beta form EGF-like domains in various
combinations. The peptides maintain the survival and proliferation of	combinations. The peptides maintain the survival and proliferation of
Schwann cells and cause proliferation, growth and differentiation of	Schwann cells and cause proliferation, growth and differentiation of
colon epithelial cells. Accordingly, they are useful to treat (in vitro	colon epithelial cells. Accordingly, they are useful to treat (in vitro
or in vivo) a disease or disorder of the colon (e.g. colitis or an	or in vivo) a disease or disorder of the colon (e.g. colitis or an
ulcer) or of the nervous system (e.g. nerve damage caused by trauma).	ulcer) or of the nervous system (e.g. nerve damage caused by trauma).
Sequence 52 AA;	Sequence 52 AA;

FT
/label= EGF_motif

FT modified_site /note= "contains 6 cysteines"
 42. .43
 FT /note= "serine-glycine dipeptide potential
 FT glycosaminoglycan addition site -
 FT not part of mature HPG-alpha sequence"
 FT 64. .65
 FT modified_site /note= "serine-glycine dipeptide potential
 FT glycosaminoglycan addition site"
 FT 151. .152
 FT modified_site /note= "serine-glycine dipeptide potential
 FT glycosaminoglycan addition site"
 FT 164. .166
 FT modified_site /note= "N-linked glycosylation site"
 FT 170. .172
 FT modified_site /note= "N-linked glycosylation site"
 FT 208. .210
 FT modified_site /note= "N-linked glycosylation site"
 FT 437. .439
 FT modified_site /note= "N-linked glycosylation site"
 FT 609. .611
 FT modified_site /note= "N-linked glycosylation site"
 FT 209. .221
 FT modified_site /note= "serine-threonine potential
 FT O-glycosylation sites"
 FT 1. .44
 FT region /note= "deduced from ORF but initiating MET is
 FT at position 45 and the processed
 FT N-terminal residue is Ser46"
 PN WO9220798-A.
 PD 26-NOV-1992.
 PF 21-MAY-1992; US-705256.
 PR 24-MAY-1991; US-765212.
 PR 25-SEP-1991; US-765212.
 PR 08-NOV-1991; US-790801.
 PR 06-MAR-1992; US-847743.
 PR 11-MAY-1992; US-880917.
 PA (GENE-) GENETECH INC.
 PI Holmes WE, Vandlen RL;
 DR WPI: 92-415776/50.
 DR P-PSDB: R29570.
 DR New polypeptide(s) heregulin(s) bind the P185HER2 receptor - for
 FT diagnosing and treating tumours and screening for agonists or
 FT antagonists for binding to P185HER2
 PS Example 3; Fig 4; 139pp; English.
 CC Heregulin-alpha was isolated from the supernatant of human breast
 CC carcinoma MDA-MB-231. After purification, a major peak of tyrosine
 CC phosphorylation activity was correlated with a 45kD protein. The
 CC protein was cleaved with lysine-C and one of the resulting peptide
 CC fragments (C-15) was used to design an oligonucleotide probe (see
 CC Q31547). The probe was labelled and used to screen an oligo
 CC dt-primed cDNA library constructed from human MDA-MB-231 cell mRNA
 CC in lambda gt10. Two positive clones, designated lambda gt10 her16
 CC and lambda gt10 her13, were identified and found to be identical.
 CC The sequence of lambda gt10 her16 contains a single ORF of 669
 CC amino acids. No stop codon was found in the translated sequence.
 CC See also Q31542-Q31548.
 SQ Sequence 669 AA;
 Query Match 41.0%; Score 113.5; DB 1; Length 669;
 Best Local Similarity 34.8%; Pred. No. 0.00017;
 Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;
 QY 1 HFKPCRDLDLAYCLNDGECFVETLTGSHKH-CRCKEGYQGVRCDO 45
 DB 222 HLVKCAEKEKFCVNGECFVWKDLSNPSRYLCKQPGFTGARTC 267
 RESULT 12
 R46918
 ID R46918 standard; Protein; 63 AA.
 AC R46918:
 DT 28-JUL-1994 (first entry)

DE EGFL2.
 KW Glial growth factor; GGF; heregulin; mitogenesis;
 KW Schwann cell; tumour; central nervous system; erbB2 receptor;
 KW antiproliferative; epidermal growth factor; EGF.
 PN WO9403644-A.
 PD 17-FEB-1994.
 PF 10-AUG-1993; U07491.
 PR 10-AUG-1992; US-927337.
 PR 25-SEP-1992; US-951747.
 PR 01-DEC-1992; US-984085.
 PR 29-JAN-1993; US-011396.
 PA (CAMP-) CAMBRIDGE NEUROSCIENCE INC.
 PI Gwynne DI, Marchionni M, McBurney RN;
 DR WPI: 94-065731/08.
 DR N-PSDB: Q58324.
 DR Glial growth factor DNA encoding numerous polypeptide factors
 FT used for inhibiting cell proliferation - for treating carcinoma
 FT and nervous disorders
 PS Disclosure; Fig 40; 178pp; English.
 CC These EGF-like domains can be required for the activation of
 CC mitogenesis in the binding reaction between GGF ligands contg.
 CC such domains and the erbB2 receptor. Pref. antiproliferative
 CC factors are those which lack these EGF-like domains.
 SQ Sequence 63 AA;
 Query Match 41.0%; Score 113.5; DB 1; Length 63;
 Best Local Similarity 34.8%; Pred. No. 2e-05; Indels 1; Gaps 1;
 Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;
 QY 1 HFKPCRDLDLAYCLNDGECFVETLTGSHKH-CRCKEGYQGVRCDO 45
 DB 2 HLVKCAEKEKFCVNGECFVWKDLSNPSRYLCKQPGFTGARTC 47
 RESULT 13
 R46921
 ID R46921 standard; Protein; 88 AA.
 AC R46921:
 DT 28-JUL-1994 (first entry)
 DE EGFL5.
 KW Glial growth factor; GGF; heregulin; mitogenesis;
 KW Schwann cell; tumour; central nervous system; erbB2 receptor;
 KW antiproliferative; epidermal growth factor; EGF.
 PN WO9403644-A.
 PD 17-FEB-1994.
 PF 10-AUG-1993; U07491.
 PR 10-AUG-1992; US-927337.
 PR 25-SEP-1992; US-951747.
 PR 01-DEC-1992; US-984085.
 PR 29-JAN-1993; US-011396.
 PA (CAMP-) CAMBRIDGE NEUROSCIENCE INC.
 PI Gwynne DI, Marchionni M, McBurney RN;
 DR WPI: 94-065731/08.
 DR N-PSDB: Q58327.
 DR Glial growth factor DNA encoding numerous polypeptide factors
 FT used for inhibiting cell proliferation - for treating carcinoma
 FT and nervous disorders
 PS Disclosure; Fig 43; 178pp; English.
 CC These GGF coding segments include regions with EGF-like homology.
 CC These EGF-like domains can be required for the activation of
 CC mitogenesis in the binding reaction between GGF ligands contg.
 CC such domains and the erbB2 receptor. Pref. antiproliferative
 CC factors are those which lack these EGF-like domains.
 SQ Sequence 88 AA;
 Query Match 41.0%; Score 113.5; DB 1; Length 88;
 Best Local Similarity 34.8%; Pred. No. 2.7e-05;
 Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;
 QY 1 HFKPCRDLDLAYCLNDGECFVETLTGSHKH-CRCKEGYQGVRCDO 45

Db 2 HLVKCAEKETFCVNGECFVWDLNPSRYLCKCQPGFTGARTC 47

RESULT 14

R46922 ID R46922 standard; Protein; 83 AA.
AC R46922; 41.0%; Score 113.5; DB 1; Length 63;
DT 28-JUL-1994 (first entry)
DE EGFL6.
KW Glial growth factor; GGF; heregulin; mitogenesis;
KW Schwann cell; tumour; central nervous system; erbB2 receptor;
KW antiproliferative; epidermal growth factor; EGF.
PN WO9403644-A.
PD 17-FEB-1994.
PF 10-AUG-1993; U07491.
PR 10-AUG-1992; US-927337.
PR 25-SEP-1992; US-951747.
PR 01-DEC-1992; US-984085.
PR 29-JAN-1993; US-011396.
PA (CAMP-) CAMBRIDGE NEUROSCIENCE INC.
PI Gwynne DI, Marchionni M, McBurney RN;
DR WPI; 94-065731/08.
DR N-PSDB; Q58328.
PT Glial growth factor DNA encoding numerous polypeptide factors
PT used for inhibiting cell proliferation - for treating carcinoma
PT and nervous disorders
PS Disclosure; Fig 44; 178pp; English.
CC The GGF coding segments include regions with EGF-like homology.
CC These EGF-like domains can be required for the activation of.
CC mitogenesis in the binding reaction between GGF ligands contg.
CC such domains and the erbB2 receptor. Pref. antiproliferative
CC factors are those which lack these EGF-like domains.
SQ Sequence 83 AA;

Query Match 41.0%; Score 113.5; DB 1; Length 83;
Best Local Similarity 34.8%; Pred. No. 2.6e-05;
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

OY 1 HFKPCRDKLAYCLNDGECFVIETLTGSHKH-CRCKEGYGVRCDO 45
Db 2 HLVKCAEKETFCVNGECFVWDLNPSRYLCKCQPGFTGARTC 47

RESULT 15

R55659 ID R55659 standard; Protein; 63 AA.
AC R55659;
DT 28-JUL-1994 (first entry)
DE EGFL2.
KW Glial growth factor; GGF; heregulin; mitogenesis;
KW Schwann cell; tumour; central nervous system;
KW epidermal growth factor; EGF.
PN WO9400140-A.
PD 06-JAN-1994.
PF 29-JUN-1993; U06228.
PR 30-JUN-1992; US-907138.
PR 03-SEP-1992; US-940389.
PR 23-OCT-1992; US-965173.
PR 24-MAR-1993; US-036555.
PA (CAMP-) CAMBRIDGE NEUROSCIENCE.
PA (LUDW-) LUDWIG INST CANCER RES.
PI Chen MS, Goodearl A, Hiles I, Marchioni M, Minghetti L;
PI Stroobant P, Waterfield M;
DR WPI; 94-025882/03.
DR N-PSDB; Q62843.
PT Glial mitogenic polypeptide factors - useful for stimulating
PT glial cell mitogenesis and treating glial cell tumours
PS Claim 53; Fig 39; 178pp; English.
CC EGFL1, EGFL2, EGFL3, EGFL4, EGFL5 and EGFL6 are used
CC for the stimulation of glial cell mitogenesis in vivo
CC and in vitro.

SQ Sequence 63 AA;

Query Match 41.0%; Score 113.5; DB 1; Length 63;
Best Local Similarity 34.8%; Pred. No. 2e-05;
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

OY 1 HFKPCRDKLAYCLNDGECFVIETLTGSHKH-CRCKEGYGVRCDO 45
Db 2 HLVKCAEKETFCVNGECFVWDLNPSRYLCKCQPGFTGARTC 47

Search completed: October 13, 1999, 22:45:05
Job time: 215 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: October 13, 1999, 22:44:43 ; Search time 15.23 Seconds
(without alignments)
189.925 Million cell updates/sec

Title: US-09-107-979-4
Perfect score: 277
Sequence: 1 HFAPCRDKLAYCLNDGECF.....SHKHCRCGEYGVRCQDFL 47

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

- Database : SPTREMBL_10.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phase.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	277	100.0	713	11	O35181 mus musculus
2	126.5	45.7	677	13	O93383 xenopus lae
3	113.5	41.0	462	4	Q12780 homo sapien
4	113.5	41.0	125	4	Q12781 homo sapien
5	113.5	41.0	640	4	Q02297 homo sapien
6	113.5	41.0	461	11	O35947 mesocricetu
7	111.5	40.3	639	11	P43323 rattus norv
8	111.5	40.3	461	11	P43325 rattus norv
9	111.5	40.3	422	11	P43326 rattus norv
10	104.5	37.7	552	4	Q12782 homo sapien
11	104.5	37.7	288	4	Q12783 homo sapien
12	104.5	37.7	175	4	Q12784 homo sapien
13	104.5	37.7	296	4	Q15491 homo sapien
14	104.5	37.7	768	4	O14667 homo sapien
15	104.5	37.7	422	4	O07111 homo sapien
16	104.5	37.7	241	4	O07110 homo sapien
17	104.5	37.7	637	4	Q02299 homo sapien
18	104.5	37.7	645	4	Q02298 homo sapien
19	104.5	37.7	241	6	O07112 bos taurus
20	102.5	37.0	662	11	P43322 rattus norv
21	102.5	37.0	304	11	P43327 rattus norv
22	102.5	37.0	230	11	P43328 rattus norv
23	102.5	37.0	636	11	P43324 rattus norv
24	98.5	35.6	685	13	O73750 gallus gall
25	98.5	35.6	677	13	O73751 gallus gall
26	98.5	35.6	480	13	O73752 gallus gall
27	98.5	35.6	602	13	Q05199 gallus gall
28	92.5	33.4	2180	5	O01768 caenorhabdi
29	90	32.5	641	11	O35073 rattus norv

ALIGNMENTS

RESULT 1

O35181 PRELIMINARY; PRT; 713 AA.
AC O35181; DT 01-JAN-1998 (TREMREL. 05, Created)
DT 01-JAN-1998 (TREMREL. 05, Last sequence update)
DT 01-MAY-1999 (TREMREL. 10, Last annotation update)
DE NEUREGULIN-3.
GN NR3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 9742072C.
RA ZHANG D., SIKKOWSKI M.X., MARK M., FRANTZ G., AKITA R., SUN Y.,
RA HILLAN K., SCROWLEY C., BRUSH J., MODOWSKI P.J.;
RT "Neuregulin-3 (NRG3): a novel neural tissue-enriched protein that
RT binds and activates ErbB4.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:9562-9567(1997).
DR EMBL; AF010130; AAB70914.1; -.
DR MGD; MGI:1097165; NR3.
DR PFAM; PF00008; EGF; 1.
KW Glycoprotein.
SQ SEQUENCE 713 AA; 77369 MW; F61BAF76 CRC32;

SEP 1997 too late

Query Match 100.0%; Score 277; DB 11; Length 713;
Best Local Similarity 100.0%; Pred. No. 2.8e-29;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HFAPCRDKLAYCLNDGECFVIETLTGSHKHCRCGEYGVRCQDFL 47
DB 288 HFAPCRDKLAYCLNDGECFVIETLTGSHKHCRCGEYGVRCQDFL 334

RESULT 2

O93383 PRELIMINARY; PRT; 677 AA.
AC O93383; DT 01-NOV-1998 (TREMREL. 08, Created)
DT 01-NOV-1998 (TREMREL. 08, Last sequence update)
DT 01-MAY-1999 (TREMREL. 10, Last annotation update)
DE NEUREGULIN ALPHA.1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Amphibia;
OC Batrachia; Anura Mesobatrachia; Pipoidae; Pipidae; Xenopodinae;
OC Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98352125.
RA YANG J.F., ZHOU R., PUN S., IP N.Y., PENG H.B., TSIM K.W.K.;

AC 035947;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
 DE NEUREGULIN.
 OS Mesocricetus auratus (Golden hamster).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; Mesocricetus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA VELASCO J.A., FEIJOO E., AVILA M.A., NOTARIO V.;
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
 DR ENBL; U96612; AAB71812.1; -;
 DR PFAM; PF00008; EGF; 1.
 DR PFAM; PF00047; ig; 1.
 KW Glycoprotein.
 SQ SEQUENCE 461 AA; 50890 MW; E937DA34 CRC32;

Query Match 41.0%; Score 113.5; DB 11; Length 461;
 Best Local Similarity 34.8%; Pred. No. 1.5e-07;
 Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;
 QY 1 HFKPCRDKLAYCLNDGECFVIELTGSKH-CRCKEGYGVRCDO 45
 DB 178 HLKCAEKETFCVNGECFVMDLSNPSRYLCKCPGFTGARCTE 223

RESULT 7
 P43323 PRELIMINARY; PRT; 639 AA.
 ID P43323
 AC P43323;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
 DE NEU DIFFERENTIATION FACTOR NDF38.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA WEN D., SUGGS S.V., KARUNAGARAN D., LIU N., CUPPLES R.L., LUO Y.,
 RA JANSSEN A.M., BEN-BARUCH N., TROLLINGER D.B., JACOBSEN V.L.,
 RA MENG S.Y., LU H.S., HU S., CHANG D., YANG W., YANIGAHARA D.,
 RA KOSKI R.A., YARDEN Y.;
 RT "Structural and functional aspects of the multiplicity of Neu
 RT differentiation factors";
 RL Mol. Cell. Biol. 14:1909-1919(1994).
 CC -/- SIMILARITY: BELONGS TO EGF FAMILY.
 DR ENBL; U02319; AAA19944.1; -;
 DR EMBL; U02323; AAA19948.1; -;
 DR PROSITE; PS00022; EGF; 1.
 DR PFAM; PF00008; EGF; 1.
 DR PFAM; PF00047; ig; 1.
 KW Transmembrane; EGF-like domain.
 FT TRANSMEM 243 265 POTENTIAL.
 SQ SEQUENCE 639 AA; 70302 MW; 1437D80A CRC32;

Query Match 40.3%; Score 111.5; DB 11; Length 639;
 Best Local Similarity 34.8%; Pred. No. 3.7e-07;
 Matches 16; Conservative 13; Mismatches 16; Indels 1; Gaps 1;
 QY 1 HFKPCRDKLAYCLNDGECFVIELTGSKH-CRCKEGYGVRCDO 45
 DB 178 HLKCAEKETFCVNGECFVMDLSNPSRYLCKCPGFTGARCTE 223

RESULT 8
 P43325 PRELIMINARY; PRT; 461 AA.
 ID P43325
 AC P43325;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
 DE NEU DIFFERENTIATION FACTOR NDF19.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA WEN D., SUGGS S.V., KARUNAGARAN D., LIU N., CUPPLES R.L., LUO Y.,
 RA JANSSEN A.M., BEN-BARUCH N., TROLLINGER D.B., JACOBSEN V.L.,
 RA MENG S.Y., LU H.S., HU S., CHANG D., YANG W., YANIGAHARA D.,
 RA KOSKI R.A., YARDEN Y.;
 RT "Structural and functional aspects of the multiplicity of Neu
 RT differentiation factors";
 RL Mol. Cell. Biol. 14:1909-1919(1994).
 CC -/- SIMILARITY: BELONGS TO EGF FAMILY.
 DR ENBL; U02316; AAA19941.1; -;
 DR EMBL; U02317; AAA19942.1; -;
 DR PROSITE; PS00022; EGF; 1.
 DR PFAM; PF00008; EGF; 1.
 DR PFAM; PF00047; ig; 1.
 KW Transmembrane; EGF-like domain.
 FT TRANSMEM 243 265 POTENTIAL.
 SQ SEQUENCE 461 AA; 50809 MW; 1289283A CRC32;

Query Match 40.3%; Score 111.5; DB 11; Length 461;
 Best Local Similarity 34.8%; Pred. No. 2.7e-07;
 Matches 16; Conservative 13; Mismatches 16; Indels 1; Gaps 1;
 QY 1 HFKPCRDKLAYCLNDGECFVIELTGSKH-CRCKEGYGVRCDO 45
 DB 178 HLKCAEKETFCVNGECFVMDLSNPSRYLCKCPGFTGARCTE 223

RESULT 9
 P43326 PRELIMINARY; PRT; 422 AA.
 ID P43326
 AC P43326;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
 DE NEU DIFFERENTIATION FACTOR NDF44.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA WEN D., SUGGS S.V., KARUNAGARAN D., LIU N., CUPPLES R.L., LUO Y.,
 RA JANSSEN A.M., BEN-BARUCH N., TROLLINGER D.B., JACOBSEN V.L.,
 RA MENG S.Y., LU H.S., HU S., CHANG D., YANG W., YANIGAHARA D.,
 RA KOSKI R.A., YARDEN Y.;
 RT "Structural and functional aspects of the multiplicity of Neu
 RT differentiation factors";
 RL Mol. Cell. Biol. 14:1909-1919(1994).
 CC -/- SIMILARITY: BELONGS TO EGF FAMILY.
 DR ENBL; M92430; -; NOT_ANNOTATED_CDS.
 DR PROSITE; PS00022; EGF; 1.
 DR PFAM; PF00008; EGF; 1.
 DR PFAM; PF00047; ig; 1.
 KW Transmembrane; EGF-like domain.

```

RESULT 11
ID Q12783 PRELIMINARY; PRT; 288 AA.
AC Q12783;
DT 01-NOV-1996 (Tremblurel. 01, Created)
DT 01-NOV-1996 (Tremblurel. 01, Last sequence update)
DT 01-NOV-1998 (Tremblurel. 08, Last annotation update)
DE NEW DIFFERENTIATION FACTOR (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
[1]
SEQUENCE FROM N.A.
RX MEDLINE; 94158863.
RA WEN D., SUGGS S.V., KARUNAGARAN D., LIU N., CUPPLES R.L., LUC
RA JANSEN A.M., BEN-BARUCH N., TROLLINGER D.B., JACOBSEN V.L.,
RA MENG S.Y., LI H.S., HU S., CHANG D., YANG W., YANIGAHARA D.,
RA KOSKI R.A., YARDEN Y.;
RT "Structural and functional aspects of the multiplicity of Neu
RT differentiation factors.";
RL Mol. Cell. Biol. 14:1909-1919(1994).

```

RESULT	13	
Q15491		
ID	Q15491	PRELIMINARY; PRT; 296 AA.
AC	Q15491;	
DT	01-NOV-1996	(TReMBLrel. 01, Created)
DT	01-NOV-1996	(TReMBLrel. 01, Last sequence update)
DT	01-NOV-1998	(TReMBLrel. 08, Last annotation update)
DE	SENSORY AND NEURON-DERIVED FACTOR.	
GN	SMDF.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;	
OC	Eutheria; Primates; Catarrhini; Homnidae; Homo.	
RP	[1]	
RN	SEQUENCE FROM N.A.	
RC	TISSUE=BRAIN STEM;	
RC	MEDLINE: 95301541;	
RA	HO W.H.; ARMANIN M.P.; NUIJENS A.; PHILLIPS H.S.; OSHEROFF P	
RT	'Sensory and motor neuron-derived factor. A novel heregulin v	
RT	highly expressed in sensory and motor neurons.";	
RL	J. Biol. Chem. 270:14523-14532(1995).	
DR	EMB1; L41827; AAC41764.1; -	

DR PFAM; PF00008; EGF; 1.
SQ SEQUENCE 296 AA; 31685 MW; 59F09FC6 CRC32;

Query Match	37.7%	Score 104.5;	DB 4;	Length 296;
Best Local Similarity	31.2%	Pred. NO. 1.5e-06;		
Matches 15;	Conservative 14;	Mismatches 18;	Indels 11;	Gaps 1;

QY 1 HFKPRDKDLAYCLNDGECVFIETLTGSKH-CRCKEGYGVRCDQL 47
 | : : : : : |||:: | : : : : : | ::
D6 233 HLVKAEKEKTFCVNGGEFGMVKDLSNPSRYICKCPNEFTGDRCNYV 280

RESULT 14

014667
ID 014667 PRELIMINARY; PRT; 768 AA.

AC	014667;	
AD	01-JAN-1998 (TRENBLrel. 05, Created)	
AE	01-JAN-1998 (TRENBLrel. 05, Last sequence update)	
AF	01-NOV-1998 (TRENBLrel. 08, Last annotation update)	
AG	GAMMA-HEREGULIN.	
AH	DE	
AI	OS	
AJ	OS	
AK	Homo sapiens (Human).	
AL	OS	
AM	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;	
AN	Eutheria; Primates; Catarrhini; Homnidae; Homo.	
AO	[1]	
AP	SEQUENCE FROM N.A.	
AQ	RP	
AR	MEDLINE: 97472144.	
AS	RA	
AT	SCHAEFER G., FITZPATRICK V.D., SLIWKOWSKI M.X.;	
AV	"Gamma-heregulin: a novel heregulin isoform that is an autocrine	
AW	growth factor for the human breast cancer cell line, MDA-MB-175."	
AX	RL	
AY	Oncoogene 15:1385-1394(1997).	
AZ	ENBL; AF009227; AAC51756.1; .	
BA	DR	
BB	PFAM; PF00008; EGF; 1.	
BC	DR	
BD	PFAM; PF00047; ig; 1.	
BE	DR	
BF	SEQUENCE 768 AA; 84177 MW; 8D5E27FF CRC32;	
BG	SQ	

Query Match 37.7%; Score 104.5; DB 4; Length 768;
Best Local Similarity 31.2%; Pred. No. 3.9e-06;
Matches 15; Conservative 14; Mismatches 18; Indels 1

QY 1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVCRDQFL 47
 | : : : : : : : : : : : : : : : :
D6 705 HLVKCAEKEKTFVNGGECEFMVKDLNSPSRYLCKCPNEFTGRCONV 752

RESULT 15

Q07111	PRELIMINARY;	PRT;	422 AA.
ID	Q07111		
AC	Q07111;		
DT	01-NOV-1996 (TREMBlrel. 01, Created)		
DT	01-NOV-1996 (TREMBlrel. 01, Last sequence update)		
DT	01-NOV-1998 (TREMBlrel. 08, Last annotation update)		
DE	GLIAL GROWTH FACTOR 2 PRECURSOR (NEUREGUIN) (GGFHP2).		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
OC	Euthera; Primates; Catarrhini; Hominoidea; Homo.		
RN	{1}		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=BRAIN;		
RX	MEDLINE: 93205115,		
RA	MARCHIONNI M.A., GOODEARL A.D., CHEN M., BIRMINGHAM-MCDONOGH		
RA	KIRK C., HENDRICKS M., DANEHY F., MISUMI D., SUDHALTER J.,		
RA	KOBAYASHI K., WRÓBLEWSKI D., LYNCH C., BALDASARRE M., HILES I		
RA	DAVIS J.B., HSUAN J., TOTTY N.F., OTSU M., MCBURNEY R.N.,		
RA	WATERFIELD M.D., STROOBANT P., GWYNNE D.;		
RT	"Glial growth factors are alternatively spliced erbB2 ligands		
RT	expressed in the nervous system.";		
RL	Nature 362:312-318(1993).		
CC	-!- FUNCTION: SCHWANN CELL MITOGEN		
CC	-!- TISSUE SPECIFICITY: EXPRESSED IN NERVOUS SYSTEM: SPINAL		
CC	CORD.		
CC	-!- SIMILARITY: BELONGS TO EGF FAMILY.		

DR	EMBL; L12260; AAR59622.1; -.
DR	PROSITE; PS00022; EGF; 1.
DR	PFAM; PF00008; EGF; 1.
DR	PFAM; PF00047; ig; 1
DR	Growth factor; EGF-like domain; Glycoprotein; Signal.
KW	Growth factor; EGF-like domain; Glycoprotein; Signal.
FT	SIGNAL 1 - 44 POTENTIAL.
FT	CHAIN 45 422 GLIAL GROWTH FACTOR 2.
FT	CHAIN 30 39 POLY-LEU.
FT	DOMAIN 101 105 POLY-ALA.
FT	DOMAIN 101 105 POLY-ALA.
FT	VARIANT 253 253 Q -> R.
FT	SEQUENCE 422 AA; 45112 MW; BC6DC7F1 CRC32;
SQ	SEQUENCE 422 AA; 45112 MW; BC6DC7F1 CRC32;

Query Match	37.7%	Score	104.5;	DB	4;	Length	422;
Best Local Similarity	31.2%;	Pred. No.	2.2e-06;				
Matches	15;	Conservative	14;	Mismatches	18;	Indels	1;
						Gaps	1;

Qy 1 HFKPCRDKDLAYCLNDGCECFVETLTGSHKH-CRCKEGYQGVRCDQFL 47
| | : | : | | | | : | : | : | : | : | :
Dd 359 HLVKCAEKEKIFCVNGGECFMVKDLSNPARYLCKCPNEFTGDRCQNYV 406

Search completed: October 14, 1999, 00:32:24
Job time: 6461 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 1999, 14:00:21 ; Search time 33.23 Seconds
(without alignments)
56.668 Million cell updates/sec

Title: US-09-107-979-4
Perfect score: 277
Sequence: 1 HFKPCRDLDLAYCLNDGECF.....SHKHCRCKEGYQGVRCDOFL 27

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR60.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	113.5	41.0	640	2	A43273	heregulin precursor
2	113.5	41.0	462	2	I38404	neu differentiation
3	113.5	41.0	125	2	I38405	neu differentiation
4	111.5	40.3	461	2	I61716	neu differentiation
5	111.5	40.3	639	2	I61719	neu differentiation
6	111.5	40.3	422	2	I61724	neu differentiation
7	105.5	38.1	125	2	S62676	heregulin isoform
8	104.5	37.7	422	2	S32357	glial growth facto
9	104.5	37.7	637	2	C43273	heregulin precursor
10	104.5	37.7	241	2	B43273	heregulin precursor
11	104.5	37.7	645	2	B43273	heregulin, splice
12	104.5	37.7	552	2	I38406	neu differentiation
13	104.5	37.7	288	2	I38407	neu differentiation
14	104.5	37.7	175	2	I38408	neu differentiation
15	104.5	37.7	296	2	A56943	sensory/motor neur
16	104.5	37.7	241	2	S32359	glial growth facto
17	102.5	37.0	636	2	I61718	neu differentiation
18	102.5	37.0	636	2	I61721	neu differentiation
19	102.5	37.0	562	2	I61722	neu differentiation
20	102.5	37.0	330	2	A56210	neu differentiation
21	102.5	37.0	304	2	I61720	neu differentiation
22	98.5	35.6	602	2	A45769	acetylcholine rece
23	90	32.5	850	2	JC5700	Erbb kinase activa
24	90	32.5	868	2	JC5701	Erbb kinase activa
25	90	32.5	860	2	JC5702	Erbb kinase activa
26	90	32.5	641	2	PC4417	Erbb kinase activa
27	85	30.7	80	1	EGVZSF	growth factor - ra
28	84.5	30.5	1220	2	A56136	jagged protein pre
29	83	30.0	162	2	S68401	epiregulin precurs
30	83	30.0	46	2	JT0747	epiregulin - rat
31	83	30.0	2531	2	A46019	gene Notch-1 prote
32	83	30.0	861	2	A48825	Notch homolog Motc
33	82.5	29.8	177	2	A37408	betacellulin precu
34	82	29.6	85	1	EGVZM1	growth factor - my
35	82	29.6	230	2	A44071	probable EGF-like
36	80.5	29.1	178	2	JCL1467	betacellulin precu
37	79.5	28.7	159	1	WFRFL1	transforming growt
38	79.5	28.7	140	1	WMVZ9	growth factor - va
39	79.5	28.7	142	1	WMVZ3C	growth factor - va

40 79.5 28.7 159 2 S27195 transforming growt
41 79.5 28.7 159 2 I57497 transforming growt
42 79 28.5 2531 2 S18188 notch protein homo
43 78 28.2 1207 1 EGHU epidermal growth f
44 77.5 28.0 722 2 I48324 DELTA-like 1 - mod
45 76 27.4 1217 1 EGSMG epidermal growth f

ALIGNMENTS

RESULT 1
A43273
heregulin precursor, splice form alpha - human
N:Alternate names: breast cancer cell differentiation factor p45; Neu differentiation
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 17-Mar-1999
C:Accession: A43273; A48498; A38155
R:Holmes, W.E.; Silwakowski, M.X.; Akita, R.W.; Henzel, W.J.; Lee, J.; Park, J.W.; Yan
Science 256, 1205-1210, 1992
A:Title: Identification of heregulin, a specific activator of p185(erbB2).
A:Reference number: A43273; MUID:92271253
A:Accession: A43273
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-640 <HOL>
A:Experimental source: breast tumor cell line, MDA-MB-231, ATCC HTB 26
A:Note: sequence extracted from NCBI backbone (NCBIP:103250)
R:Culouscou, J.M.; Plozman, G.D.; Carlton, G.W.; Green, J.M.; Shoyab, M.
J. Biol. Chem. 268, 18407-18410, 1993
A:Title: Characterization of a breast cancer cell differentiation factor that specifi
A:Reference number: A48498; MUID:93366731
A:Accession: A48498
A:Molecule type: protein
A:Residues: 20-21, 'X', 23-24, 'XX', 27-28 <CUL>
R:Peles, E.; Bacus, S.S.; Koski, R.A.; Lu, H.S.; Wen, D.; Ogden, S.G.; Levy, R.B.; Ya
Cell 69, 205-216, 1992
A:Title: Isolation of the neu/HER-2 stimulatory ligand: a 44 kd glycoprotein that ind
A:Reference number: A38155; MUID:92208945
A:Accession: A38155
A:Molecule type: protein
A:Residues: 'X', 13-16, 'X', 18-20, 'RG', 23-24, 'GP', 27, 'E', 29, 'XP', 32-36 <PEL>
A:Note: sequence extracted from NCBI backbone (NCBIP:91347)
C:Genetics:
A:Gene: GDB: HGL
A:Cross-references: GDB:132656; OMIM:142445
A:Map position: 8p22-8p11
C:Superfamily: EGF homology
C:Keywords: alternative splicing; glycoprotein
F:182-221/Domain: EGF homology <EGF>

Query Match 41.0%; Score 113.5; DB 2; Length 640;
Best Local Similarity 34.8%; Pred. No. 5.6e-06;
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

Qy 1 HFKPCRDLDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDO 45
Db 178 HLVCRAEKRETCVNGGECFMVKDLNSPRLCKCQPGTGARCTE 223

RESULT 2
I38404
neu differentiation factor - human
C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 29-May-1998
C:Accession: I38404
R:Wen, D.; Suggs, S.V.; Karunakaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.
Mol. Cell. Biol. 14, 1909-1919, 1994
A:Title: Structural and functional aspects of the multiplicity of Neu differentiation
A:Reference number: A56210; MUID:94158863
A:Accession: I38404
A:Status: preliminary; translated from GB/EMBL/DBJ

```

A:Molecule type: mRNA
A:Residues: 1-462 <RES>
A:Cross-references: EMBL:U02326; NID:g408402; PID:g408403

Query Match      41.0%; Score 113.5; DB 2; Length 462;
Best Local Similarity 34.8%; Pred. No. 4.2e-06;
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

Qy    1 HFKPCRDRLAYCLNDGECFVIETLTGSHKH-CRCCKEGYQGVRCDQ 45
      | :|: |: ||||::| ::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
DBD   178 HLVKCAEKEKTCVNGGCFWVKDLNPSRYLCRKCQPFGTGARCTE 223

RESULT 3
I38405
neu differentiation factor - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 29-May-1998
C:Accession: I38405
C:RWen, D.; Suggs, S.V.; Karunakaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.M.; Mol. Cell. Biol. 14, 1909-1919, 1994
A>Title: Structural and functional aspects of the multiplicity of Neu differentiation fa
A:Reference number: A56210; MUID:94158863
A:Accession: I38405
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-125 <RES>
A:Cross-references: EMBL:U02327; NID:g408404; PID:g408405

Query Match      41.0%; Score 113.5; DB 2; Length 125;
Best Local Similarity 34.8%; Pred. No. 1.3e-06;
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

Qy    1 HFKPCRDKLAYCLNDGECFVIETLTGSHKH-CRCCKEGYQGVRCDQ 45
      | :|: |: ||||::| ::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
DBD   56 HLVKCAEKEKTCVNGGCFWVKDLNPSRYLCRKCQPFGTGARCTE 101

RESULT 4
I61716
neu differentiation factor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 03-Aug-1998
C:Accession: I61716; I61717
C:RWen, D.; Suggs, S.V.; Karunakaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.M.; Mol. Cell. Biol. 14, 1909-1919, 1994
A>Title: Structural and functional aspects of the multiplicity of Neu differentiation fa
A:Reference number: A56210; MUID:94158863
A:Accession: I61716
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-461 <RES>
A:Cross-references: EMBL:U02316; NID:g408382; PID:g408383
A:Accession: I61717
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-461 <RE2>
A:Cross-references: EMBL:U02317; NID:g408384; PID:g408385

Query Match      40.3%; Score 111.5; DB 2; Length 461;
Best Local Similarity 34.8%; Pred. No. 7.1e-06;
Matches 16; Conservative 13; Mismatches 16; Indels 1; Gaps 1;

Qy    1 HFKPCRDKLAYCLNDGECFVIETLTGSHKH-CRCCKEGYQGVRCDQ 45
      | :|: |: ||||::| ::|:~
DBD   178 HLKCAEKEKTCVNGGCFWVKDLNPSRYLCRKCQPFGTGARCTE 223

RESULT 5
```

A:Title: Isolation and structural characterization of recombinant human neu differentiation
 A:Reference number: S62676; MUID:96139341
 A:Accession: S62676
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-67-16;17-30;31-38;39-58;59-92;93-120;121-125 <HAR>
 C:Keywords: proto-oncogene

Query Match 38.1%; Score 105.5; DB 2; Length 125;
 Best Local Similarity 32.6%; Pred. No. 1.1e-05;
 Matches 15; Conservative 14; Mismatches 16; Indels 1; Gaps 1;

QY 1 HFKPCRDKDLAYCLNDGECFVITLTGSHKH-CRCRKGQGVRCDOFL 45
 DB 75 HLKCAEKEKTCFVNGGECFMVKDLSNPSRYLCKCPQFTGRCCTE 120

RESULT 8

glial growth factor - human
 C:Species: Homo sapiens (man)
 C:Date: 02-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 17-Mar-1999
 A:Accession: S32357
 R:Marchionni, M.A.; Goodearl, A.D.J.; Chen, M.S.; Bermingham-McDonogh, O.; Kirk, C.; Hen-
 les, I.; Davis, J.B.; Hsuan, J.J.; Totty, N.F.; Otsu, M.; McBurney, R.N.; Waterfield, M.
 Nature 362, 312-318, 1993
 A:Title: Glial growth factors are alternatively spliced erbB2 ligands expressed in the n
 A:Reference number: S32357; MUID:93205115
 A:Accession: S32357
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-422 <MAR>
 C:Superfamily: EGF homology
 F:363-402/Domain: EGF homology <EGF>

Query Match 37.7%; Score 104.5; DB 2; Length 422;
 Best Local Similarity 31.2%; Pred. No. 4.3e-05;
 Matches 15; Conservative 14; Mismatches 18; Indels 1; Gaps 1;

QY 1 HFKPCRDKDLAYCLNDGECFVITLTGSHKH-CRCRKGQGVRCDOFL 47
 DB 359 HLKCAEKEKTCFVNGGECFMVKDLSNPSRYLCKCPNEFTGRCQNYV 406

RESULT 9

heregulin precursor, splice form beta-2 - human
 C:Species: Homo sapiens (man)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 14-Aug-1998
 A:Accession: C43273
 R:Holmes, W.E.; Sliwkowski, M.X.; Akita, R.W.; Henzel, W.J.; Lee, J.; Park, J.W.; Yansu
 Science 256, 1205-1210, 1992
 A:Title: Identification of heregulin, a specific activator of p185(erbB2).
 A:Reference number: A43273; MUID:92271253
 A:Accession: C43273
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
 A:Molecule type: mRNA
 A:Residues: 1-637 <HOL>
 C:Genetics:
 A:Gene: GDB:HGL
 A:Cross-references: GDB:132656; OMIM:142445
 A:Map position: 8p22-8p11
 C:Superfamily: EGF homology
 C:Keywords: alternative splicing
 F:182-221/Domain: EGF homology <EGF>

Query Match 37.7%; Score 104.5; DB 2; Length 637;
 Best Local Similarity 31.2%; Pred. No. 6.2e-05;
 Matches 15; Conservative 14; Mismatches 18; Indels 1; Gaps 1;

QY 1 HFKPCRDKDLAYCLNDGECFVITLTGSHKH-CRCRKGQGVRCDOFL 47
 DB 178 HLKCAEKEKTCFVNGGECFMVKDLSNPSRYLCKCPNEFTGRCQNYV 225

RESULT 10

heregulin precursor, splice form beta-3 - human
 N:Alternate names: glial growth factor HRG-beta-3; heregulin
 C:Species: Homo sapiens (man)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 17-Mar-1999
 A:Accession: D43273; S32358
 R:Holmes, W.E.; Sliwkowski, M.X.; Akita, R.W.; Henzel, W.J.; Lee, J.; Park, J.W.; Yan
 Science 256, 1205-1210, 1992
 A:Title: Identification of heregulin, a specific activator of p185(erbB2).
 A:Reference number: A43273; MUID:92271253
 A:Accession: D43273
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual

A:Molecule type: mRNA
 A:Residues: 1-241 <HOL>
 R:Marchionni, M.A.; Goodearl, A.D.J.; Chen, M.S.; Bermingham-McDonogh, O.; Kirk, C.;
 les, I.; Davis, J.B.; Hsuan, J.J.; Totty, N.F.; Otsu, M.; McBurney, R.N.; Waterfield,
 Nature 362, 312-318, 1993
 A:Title: Glial growth factors are alternatively spliced erbB2 ligands expressed in th
 A:Reference number: S32357; MUID:93205115
 A:Accession: S32358
 A:Molecule type: mRNA
 A:Residues: 1-241 <MAR>
 C:Genetics:
 A:Gene: GDB:HGL; GGF
 A:Cross-references: GDB:132656; OMIM:142445
 A:Map position: 8p22-8p11
 C:Superfamily: EGF homology
 C:Keywords: alternative splicing
 F:182-221/Domain: EGF homology <EGF>

Query Match 37.7%; Score 104.5; DB 2; Length 241;
 Best Local Similarity 31.2%; Pred. No. 2.7e-05;
 Matches 15; Conservative 14; Mismatches 18; Indels 1; Gaps 1;

QY 1 HFKPCRDKDLAYCLNDGECFVITLTGSHKH-CRCRKGQGVRCDOFL 47
 DB 178 HLKCAEKEKTCFVNGGECFMVKDLSNPSRYLCKCPNEFTGRCQNYV 225

RESULT 11

heregulin, splice form beta 1 - human
 C:Species: Homo sapiens (man)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 14-Aug-1998
 A:Accession: B43273
 R:Holmes, W.E.; Sliwkowski, M.X.; Akita, R.W.; Henzel, W.J.; Lee, J.; Park, J.W.; Yan
 Science 256, 1205-1210, 1992
 A:Title: Identification of heregulin, a specific activator of p185(erbB2).
 A:Reference number: A43273; MUID:92271253
 A:Accession: B43273
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual

A:Molecule type: mRNA
 A:Residues: 1-645 <HOL>
 C:Genetics:
 A:Gene: GDB:HGL
 A:Cross-references: GDB:132656; OMIM:142445
 A:Map position: 8p22-8p11
 C:Superfamily: EGF homology
 C:Keywords: alternative splicing
 F:182-221/Domain: EGF homology <EGF>

Query Match 37.7%; Score 104.5; DB 2; Length 645;
 Best Local Similarity 31.2%; Pred. No. 6.3e-05;
 Matches 15; Conservative 14; Mismatches 18; Indels 1; Gaps 1;

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 1999, 01:54:24 ; Search time 10.23 Seconds
(without alignments)
129.874 Million cell updates/sec

Title: US-09-107-979-4
Perfect score: 277
Sequence: 1 HFPCRKDLAYCLNDGEF.....SHKCRCKEGYQVRCQDFL 47

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt_37.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	32.5	850	1	NTAK_HUMAN
2	90	32.5	868	1	NTAK_RAT
3	85	30.7	80	1	GRFA_SFVKA
4	83	30.0	2531	1	NTC1_MOUSE
5	82.5	29.8	177	1	BTC_MOUSE
6	82	29.6	85	1	GRFA_MXVL
7	82	29.6	230	1	SPIT_DROME
8	80.5	29.1	178	1	BTC_HUMAN
9	79.5	28.7	142	1	GRFA_VACCC
10	79.5	28.7	140	1	GRFA_VACCV
11	79.5	28.7	159	1	TGFA_MOUSE
12	79.5	28.7	159	1	TGFA_RAT
13	79	28.5	2531	1	NTC1_RAT
14	78	28.2	1207	1	EGF_HUMAN
15	77.5	28.0	722	1	DLI1_MOUSE
16	77.5	28.0	714	1	DLI1_RAT
17	76	27.4	2139	1	CRB_DROME
18	76	27.4	1217	1	EGF_MOUSE
19	76	27.4	484	1	LEM2_PIG
20	75	27.1	611	1	LEM2_CANFA
21	74.5	26.9	723	1	DLI1_HUMAN
22	74.5	26.9	603	1	FAI2_CAVPO
23	74.5	26.9	1429	1	LI12_CAEEL
24	74	26.7	2871	1	FBNI_BOVIN
25	74	26.7	2871	1	FBNI_HUMAN
26	74	26.7	2871	1	FBNI_MOUSE
27	74	26.7	610	1	LEM2_HUMAN
28	74	26.7	551	1	LEM2_RABIT
29	73	26.4	294	1	GRK_DROME
30	73	26.4	409	1	MFGM_PIG
31	73	26.4	2524	1	NOTC_XENLA
32	73	26.4	1964	1	NTCA_MOUSE
33	72.5	26.2	615	1	FAI2_HUMAN
34	72.5	26.2	208	1	HGFE_MOUSE
35	72.5	26.2	463	1	MFGM_MOUSE
36	72.5	26.2	121	1	TGFA_MACMO
37	72	26.0	1133	1	EGF_RAT
38	72	26.0	570	1	FBP3_STRPU
39	72	26.0	768	1	LEM3_MOUSE
40	71.5	25.8	4543	1	LRP1_CHICK
41	71.5	25.8	160	1	TGFA_HUMAN
42	71.5	25.8	50	1	TGFA_RABIT
43	71	25.6	53	1	EGF_PIG

RESULT 1	NTAK_HUMAN	STANDARD:	PRT:	850 AA.
ID	NTAK_HUMAN			
AC	014511;			
DT	15-DEC-1998 (REL. 37, CREATED)			
DT	15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)			
DT	15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)			
DE	NTAK PROTEIN (NEURAL- AND THYMUS- DERIVED ACTIVATOR FOR ERBB KINASES).			
GN	NTAK.			
OS	HOMO SAPIENS (HUMAN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;			
OC	PRIMATES; CATARRHINI; HOMINIDAE; HOMO.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=NEUROBLASTOMA;			
RX	MEDLINE: 98006324			
RA	HIGASHIYAMA S., HORIKAWA M., YAMADA K., ICHINO N., NAKANO N.,			
RA	NAKAGAWA T., MIYAGAWA J., MATSUSHITA N., NAGATSU T., TANIGUCHI N.,			
RA	ISHIGURO H.;			
RT	"A novel brain-derived member of the epidermal growth factor family			
RT	that interacts with ErbB3 and ErbB4.";			
RL	J. BIOCHEM. 122:675-680(1997).			
CC	-!- FUNCTION: LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE RECEPTORS.			
CC	-!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.			
CC	-!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS			
CC	1 C2-LIKE DOMAIN.			
CC	-!- SIMILARITY: BELONGS TO EGF FAMILY. TO HEPARIN-BINDING PROTEINS.			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	EMBL; AB005060; J1024296; -			
DR	PROSITE; PS00022; EGF_1; 1.			
DR	PROSITE; PS01186; EGF_2; 1.			
DR	PFAM; PF00008; EGF; 1.			
DR	PFAM; PF00047; 19; 1.			
KW	TRANSMEMBRANE; EGF-LIKE DOMAIN; IMMUNOGLOBULIN FOLD; GLYCOPROTEIN.			
FT	TRANSMEM 406 426			
FT	POTENTIAL.			
FT	IG-LIKE C2-TYPE DOMAIN.			
FT	DOMAIN 250 318			
FT	EGF-LIKE.			
FT	DOMAIN 341 382			
FT	BY SIMILARITY.			
FT	DISULFID 257 311			
FT	BY SIMILARITY.			
FT	DISULFID 345 359			
FT	BY SIMILARITY.			
FT	DISULFID 372 381			
FT	BY SIMILARITY.			
FT	DOMAIN 10 13			
FT	POLY-PRO.			
FT	DOMAIN 20 30			
FT	POLY-SER.			
FT	DOMAIN 33 47			
FT	POLY-SER.			
FT	DOMAIN 87 90			
FT	POLY-ALA.			
FT	DOMAIN 721 727			
FT	POLY-PRO.			
FT	CARBOHYD 52 52			
FT	POTENTIAL.			
FT	CARBOHYD 53 53			
FT	POTENTIAL.			
FT	CARBOHYD 147 147			
FT	POTENTIAL.			
FT	CARBOHYD 278 278			
FT	POTENTIAL.			
FT	CARBOHYD 346 346			
FT	POTENTIAL.			
SQ	SEQUENCE 850 AA; 91678 MW; DES08427 CRC32;			

Query Match 32.5%; Score 90; DB 1; Length 850;
Best Local Similarity 34.9%; Pred. No. 0.0016;
Matches 15; Conservative 9; Mismatches 17; Indels 2; Gaps 1;

OY	1	HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYGVRC	43	
Db	341	HARKNETAKSYCVNGVCYIEGI--NOLSKCPNGFFGQRC	381	
RESULT	2			
NTAK_RAT				
ID	NTAK_RAT	STANDARD; PRT;	868 AA.	
AC	O35569; O35570; O35571; O35572;			
DT	15-DEC-1998 (REL. 37, CREATED)			
DT	15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)			
DT	15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)			
DE	NTAK PROTEIN PRECURSOR (NEURAL- AND THYMUS- DERIVED ACTIVATOR FOR ERBB KINASES).			
DE	NTAK.			
GN	RATTUS NORVEGICUS (RAT).			
OS	RATTUS NORVEGICUS (RAT).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;			
OC	RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.			
RN	[1]			
RP	SEQUENCE FROM N.A., SEQUENCE OF 128-162, AND ALTERNATIVE SPLICING.			
RX	MEDLINE; 98006324.			
RA	HIGASHIYAMA S., HORIKAWA M., YAMADA K., ICHINO N., NAKANO N.,			
RA	NAGAGAWA T., MIYAGAWA J., MATSUSHITA N., NAGATSU T., TANIGUCHI N.,			
RA	ISHIGURO H.;			
RT	"A novel brain-derived member of the epidermal growth factor family that interacts with ErbB3 and ErbB4.";			
RL	J. BIOCHEM. 122:675-680(1997).			
CC	-!- FUNCTION: LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE RECEPTORS.			
CC	-!- DEVELOPMENTAL STAGE: EXPRESSED ONLY IN THE BRAIN OF E11.5 EMBRYOS			
CC	AND IN THE BRAIN (ESPECIALLY THE OLFACTORY BULB AND CEREBELLUM)			
CC	AND THYMUS OF ADULT RATS.			
CC	-!- ALTERNATIVE PRODUCTS: AT LEAST FIVE ISOFORMS; NTA-ALPHA1 (SHOWN HERE), NTA-ALPHA2A, NTA-ALPHA2B, NTA-ALPHA2C, AND NTA-ALPHA2D ARE PRODUCED BY ALTERNATIVE SPLICING.			
CC	-!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.			
CC	-!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS 1 C2-LIKE DOMAIN.			
CC	-!- SIMILARITY: BELONGS TO EGF FAMILY. TO HEPARIN-BINDING PROTEINS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; D89995; D1024223;			
DR	EMBL; D89996; D1024224;			
DR	EMBL; D89997; D1024225;			
DR	EMBL; D89998; D1024226;			
DR	PROSITE; PS00022; EGF_1; 1.			
DR	PROSITE; PS01186; EGF_2; 1.			
DR	PFAM; PF00008; EGF; 1.			
DR	PFAM; PF00047; ig; 1.			
KW	TRANSMEMBRANE; EGF-LIKE DOMAIN; IMMUNOGLOBULIN FOLD; GLYCOPROTEIN;			
KW	ALTERNATIVE SPLICING			
FT	PROPEP	1	127	POTENTIAL.
FT	CHAIN	128	868	NTAK PROTEIN.
FT	TRANSMEM	430	450	POTENTIAL.
FT	DOMAIN	266	334	IG-LIKE C2-TYPE DOMAIN.
FT	DOMAIN	357	398	EGF-LIKE.
FT	DISULFID	273	327	BY SIMILARITY.
FT	DISULFID	361	375	BY SIMILARITY.
FT	DISULFID	369	386	BY SIMILARITY.
FT	DISULFID	388	397	BY SIMILARITY.
FT	DOMAIN	22	32	POLY-SER.
FT	DOMAIN	35	45	POLY-SER.
FT	DOMAIN	56	59	POLY-THR.
FT	DOMAIN	103	106	POLY-ALA.
FT	DOMAIN	739	745	POLY-PRO.
FT	CARBOHYD	33	33	POTENTIAL.
FT	CARBOHYD	34	34	POTENTIAL.

Query Match

Best Local Similarity

Matches

Score 85;

DB 1;

Length 80;

Pred. No. 0.00077;

Mismatches 17;

Indels 8;

Gaps 3;

OY	1	HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYGVRC	47	
Db	29	HYVCNHDYENYCLANGTCFTIADLNVST---PFCVKRINTEGSRG-QFI	75	

Query Match

Best Local Similarity

Matches

Score 85;

DB 1;

Length 80;

Pred. No. 0.00077;

Mismatches 17;

Indels 8;

Gaps 3;

OY	1	HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYGVRC	43	
Db	357	HARKNETAKSYCVNGVCYIEGI--NOLSKCPNGFFGQRC	397	

Query Match

Best Local Similarity

Matches

Score 85;

DB 1;

Length 868;

Pred. No. 0.0017;

Mismatches 15;

Indels 2;

Gaps 1;

FT	CARBOHYD	163	POTENTIAL.
FT	CARBOHYD	294	POTENTIAL.
FT	CARBOHYD	362	POTENTIAL.
FT	VARSPLIC	388	C->G (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	389	MISSING (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	390	NGFFQRCLEKFLRLYMPDPKQKHLGELKE--> VGYTG
FT	VARSPLIC	421	DRCOQFAMVNFSK (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	414	MISSING (IN ISOFORM NTA-ALPHA2).
FT	SEQUENCE	868 AA;	C202303C CRC32;
SQ	SEQUENCE	868 AA;	93776 MW;

Query Match

Best Local Similarity

Matches

Score 85;

DB 1;

Length 868;

Pred. No. 0.0017;

Mismatches 15;

Indels 2;

Gaps 1;

OY	1	HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYGVRC	43	
Db	357	HARKNETAKSYCVNGVCYIEGI--NOLSKCPNGFFGQRC	397	

Query Match

Best Local Similarity

Matches

Score 85;

DB 1;

Length 868;

Pred. No. 0.0017;

Mismatches 15;

Indels 2;

Gaps 1;

FT	CARBOHYD	163	POTENTIAL.
FT	CARBOHYD	294	POTENTIAL.
FT	CARBOHYD	362	POTENTIAL.
FT	VARSPLIC	388	C->G (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	389	MISSING (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	390	NGFFQRCLEKFLRLYMPDPKQKHLGELKE--> VGYTG
FT	VARSPLIC	421	DRCOQFAMVNFSK (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	414	MISSING (IN ISOFORM NTA-ALPHA2).
FT	SEQUENCE	868 AA;	C202303C CRC32;
SQ	SEQUENCE	868 AA;	93776 MW;

Query Match

Best Local Similarity

Matches

Score 85;

DB 1;

Length 868;

Pred. No. 0.0017;

Mismatches 15;

Indels 2;

Gaps 1;

OY	1	HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYGVRC	43	
Db	357	HARKNETAKSYCVNGVCYIEGI--NOLSKCPNGFFGQRC	397	

Query Match

Best Local Similarity

Matches

Score 85;

DB 1;

Length 868;

Pred. No. 0.0017;

Mismatches 15;

Indels 2;

Gaps 1;

FT	CARBOHYD	163	POTENTIAL.
FT	CARBOHYD	294	POTENTIAL.
FT	CARBOHYD	362	POTENTIAL.
FT	VARSPLIC	388	C->G (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	389	MISSING (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	390	NGFFQRCLEKFLRLYMPDPKQKHLGELKE--> VGYTG
FT	VARSPLIC	421	DRCOQFAMVNFSK (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	414	MISSING (IN ISOFORM NTA-ALPHA2).
FT	SEQUENCE	868 AA;	C202303C CRC32;
SQ	SEQUENCE	868 AA;	93776 MW;

Query Match

Best Local Similarity

Matches

Score 85;

DB 1;

Length 868;

Pred. No. 0.0017;

Mismatches 15;

Indels 2;

Gaps 1;

OY	1	HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYGVRC	43	
Db	357	HARKNETAKSYCVNGVCYIEGI--NOLSKCPNGFFGQRC	397	

Query Match

Best Local Similarity

Matches

Score 85;

DB 1;

Length 868;

Pred. No. 0.0017;

Mismatches 15;

Indels 2;

Gaps 1;

FT	CARBOHYD	163	POTENTIAL.
FT	CARBOHYD	294	POTENTIAL.
FT	CARBOHYD	362	POTENTIAL.
FT	VARSPLIC	388	C->G (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	389	MISSING (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	390	NGFFQRCLEKFLRLYMPDPKQKHLGELKE--> VGYTG
FT	VARSPLIC	421	DRCOQFAMVNFSK (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	414	MISSING (IN ISOFORM NTA-ALPHA2).
FT	SEQUENCE	868 AA;	C202303C CRC32;
SQ	SEQUENCE	868 AA;	93776 MW;

Query Match

Best Local Similarity

Matches

Score 85;

DB 1;

Length 868;

Pred. No. 0.0017;

Mismatches 15;

Indels 2;

Gaps 1;

OY	1	HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYGVRC	43	
Db	357	HARKNETAKSYCVNGVCYIEGI--NOLSKCPNGFFGQRC	397	

Query Match

Best Local Similarity

Matches

Score 85;

DB 1;

Length 868;

Pred. No. 0.0017;

Mismatches 15;

Indels 2;

Gaps 1;

FT	CARBOHYD	163	POTENTIAL.
FT	CARBOHYD	294	POTENTIAL.
FT	CARBOHYD	362	POTENTIAL.
FT	VARSPLIC	388	C->G (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	389	MISSING (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	390	NGFFQRCLEKFLRLYMPDPKQKHLGELKE--> VGYTG
FT	VARSPLIC	421	DRCOQFAMVNFSK (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	414	MISSING (IN ISOFORM NTA-ALPHA2).
FT	SEQUENCE	868 AA;	C202303C CRC32;
SQ	SEQUENCE	868 AA;	93776 MW;

Query Match

Best Local Similarity

Matches

Score 85;

DB 1;

Length 868;

Pred. No. 0.0017;

Mismatches 15;

Indels 2;

Gaps 1;

OY	1	HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYGVRC	43	
Db	357	HARKNETAKSYCVNGVCYIEGI--NOLSKCPNGFFGQRC	397	

Query Match

Best Local Similarity

Matches

Score 85;

DB 1;

Length 868;

Pred. No. 0.0017;

Mismatches 15;

Indels 2;

Gaps 1;

FT	CARBOHYD	163	POTENTIAL.
FT	CARBOHYD	294	POTENTIAL.
FT	CARBOHYD	362	POTENTIAL.
FT	VARSPLIC	388	C->G (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	389	MISSING (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	390	NGFFQRCLEKFLRLYMPDPKQKHLGELKE--> VGYTG
FT	VARSPLIC	421	DRCOQFAMVNFSK (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	414	MISSING (IN ISOFORM NTA-ALPHA2).
FT	SEQUENCE	868 AA;	C202303C CRC32;
SQ	SEQUENCE	868 AA;	93776 MW;

Query Match

Best Local Similarity

Matches

Score 85;

DB 1;

Length 868;

Pred. No. 0.0017;

Mismatches 15;

Indels 2;

Gaps 1;

OY	1	HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYGVRC	43	
Db	357	HARKNETAKSYCVNGVCYIEGI--NOLSKCPNGFFGQRC	397	

Query Match

Best Local Similarity

Matches

Score 85;

DB 1;

Length 868;

Pred. No. 0.0017;

Mismatches 15;

Indels 2;

Gaps 1;

FT	CARBOHYD	163	POTENTIAL.
FT	CARBOHYD	294	POTENTIAL.
FT	CARBOHYD	362	POTENTIAL.
FT	VARSPLIC	388	C->G (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	389	MISSING (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	390	NGFFQRCLEKFLRLYMPDPKQKHLGELKE--> VGYTG
FT	VARSPLIC	421	DRCOQFAMVNFSK (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	414	MISSING (IN ISOFORM NTA-ALPHA2).
FT	SEQUENCE	868 AA;	C202303C CRC32;
SQ	SEQUENCE	868 AA;	93776 MW;

Query Match

Best Local Similarity

Matches

Score 85;

DB 1;

Length 868;

Pred. No. 0.0017;

Mismatches 15;

Indels 2;

Gaps 1;

OY	1	HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYGVRC	43	
Db	357	HARKNETAKSYCVNGVCYIEGI--NOLSKCPNGFFGQRC	397	

Query Match

Best Local Similarity

Matches

Score 85;

DB 1;

Length 868;

Pred. No. 0.0017;

Mismatches 15;

Indels 2;

Gaps 1;

FT	CARBOHYD	163	POTENTIAL.
FT	CARBOHYD	294	POTENTIAL.
FT	CARBOHYD	362	POTENTIAL.
FT	VARSPLIC	388	C->G (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	389	MISSING (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	390	NGFFQRCLEKFLRLYMPDPKQKHLGELKE--> VGYTG
FT	VARSPLIC	421	DRCOQFAMVNFSK (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	414	MISSING (IN ISOFORM NTA-ALPHA2).
FT	SEQUENCE	868 AA;	C202303C CRC32;
SQ	SEQUENCE	868 AA;	93776 MW;

Query Match

Best Local Similarity

Matches

Score 85;

DB 1;

Length 868;

Pred. No. 0.0017;

Mismatches 15;

Indels 2;

Gaps 1;

OY	1	HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYGVRC	43	
Db	357	HARKNETAKSYCVNGVCYIEGI--NOLSKCPNGFFGQRC	397	

Query Match

Best Local Similarity

Matches

Score 85;

DB 1;

Length 868;

Pred. No. 0.0017;

Mismatches 15;

Indels 2;

Gaps 1;

FT	CARBOHYD	163	POTENTIAL.
FT	CARBOHYD	294	POTENTIAL.
FT	CARBOHYD	362	POTENTIAL.
FT	VARSPLIC	388	C->G (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	389	MISSING (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	390	NGFFQRCLEKFLRLYMPDPKQKHLGELKE--> VGYTG
FT	VARSPLIC	421	DRCOQFAMVNFSK (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	414	MISSING (IN ISOFORM NTA-ALPHA2).
FT	SEQUENCE	868 AA;	C202303C CRC32;
SQ	SEQUENCE	868 AA;	93776 MW;

Query Match

Best Local Similarity

Matches

Score 85;

DB 1;

Length 868;

Pred. No. 0.0017;

Mismatches 15;

Indels 2;

Gaps 1;

OY	1	HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYGVRC	43	
Db	357	HARKNETAKSYCVNGVCYIEGI--NOLSKCPNGFFGQRC	397	

Query Match

Best Local Similarity

Matches

Score 85;

DB 1;

Length 868;

Pred. No. 0.0017;

Mismatches 15;

Indels 2;

Gaps 1;

FT	CARBOHYD	163	POTENTIAL.
FT	CARBOHYD	294	POTENTIAL.
FT	CARBOHYD	362	POTENTIAL.
FT	VARSPLIC	388	C->G (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	389	MISSING (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	390	NGFFQRCLEKFLRLYMPDPKQKHLGELKE--> VGYTG
FT	VARSPLIC	421	DRCOQFAMVNFSK (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	414	MISSING (IN ISOFORM NTA-ALPHA2).
FT	SEQUENCE	868 AA;	C202303C CRC32;
SQ	SEQUENCE	868 AA;	93776 MW;

Query Match

Best Local Similarity

Matches

Score 85;

DB 1;

Length 868;

Pred. No. 0.0017;

Mismatches 15;

Indels 2;

Gaps 1;

OY	1	HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYGVRC	43	
Db	357	HARKNETAKSYCVNGVCYIEGI--NOLSKCPNGFFGQRC	397	

Query Match

Best Local Similarity

Matches

Score 85;

DB 1;

Length 868;

Pred. No. 0.0017;

Mismatches 15;

Indels 2;

Gaps 1;

FT	CARBOHYD	163	POTENTIAL.
FT	CARBOHYD	294	POTENTIAL.
FT	CARBOHYD	362	POTENTIAL.
FT	VARSPLIC	388	C->G (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	389	MISSING (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	390	NGFFQRCLEKFLRLYMPDPKQKHLGELKE--> VGYTG
FT	VARSPLIC	421	DRCOQFAMVNFSK (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	414	MISSING (IN ISOFORM NTA-ALPHA2).
FT	SEQUENCE	868 AA;	C202303C CRC32;
SQ	SEQUENCE	868 AA;	93776 MW;

Query Match

Best Local Similarity

Matches

Score 85;

DB 1;

Length 868;

Pred. No. 0.0017;

Mismatches 15;

Indels 2;

Gaps 1;

OY	1	HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYGVRC	43	
Db	357	HARKNETAKSYCVNGVCYIEGI--NOLSKCPNGFFGQRC	397	

Query Match

Best Local Similarity

Matches

Score 85;

DB 1;

Length 868;

Pred. No. 0.0017;

Mismatches 15;

Indels 2;

Gaps 1;

FT	CARBOHYD	163	POTENTIAL.
FT	CARBOHYD	294	POTENTIAL.
FT	CARBOHYD	362	POTENTIAL.
FT	VARSPLIC	388	C->G (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	389	MISSING (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	390	NGFFQRCLEKFLRLYMPDPKQKHLGELKE--> VGYTG
FT	VARSPLIC	421	DRCOQFAMVNFSK (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	414	MISSING (IN ISOFORM NTA-ALPHA2).
FT	SEQUENCE	868 AA;	C202303C CRC32;
SQ	SEQUENCE	868 AA;	93776 MW;

Query Match

Best Local Similarity

Matches

Score 85;

DB 1;

Length 868;

Pred. No. 0.0017;

Mismatches 15;

Indels 2;

Gaps 1;

OY	1	HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYGVRC	43	
Db	357	HARKNETAKSYCVNGVCYIEGI--NOLSKCPNGFFGQRC	397	

Query Match

Best Local Similarity

Matches

Score 85;

DB 1;

Length 868;

Pred. No. 0.0017;

Mismatches 15;

Indels 2;

Gaps 1;

FT	CARBOHYD	163	POTENTIAL.
FT	CARBOHYD	294	POTENTIAL.
FT	CARBOHYD	362	POTENTIAL.
FT	VARSPLIC	388	C->G (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	389	MISSING (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	390	NGFFQRCLEKFLRLYMPDPKQKHLGELKE--> VGYTG
FT	VARSPLIC	421	DRCOQFAMVNFSK (IN ISOFORM NTA-ALPHA2).
FT	VARSPLIC	414	MISSING (IN ISOFORM

```
RESULT 4
NTCL_MOUSE STANDARD: PRT; 2531 AA.
AC Q01705;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1 PRECURSOR (NOTCH PROTEIN).
GN NOTCH1 OR NOTCH.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
[1]
SEQUENCE FROM N.A.
RN RP SEQUENCE OF 1551-2170 FROM N.A.
RC TISSUE=EMBRYO;
RX MEDLINE; 93048835.
RA FRANCO DEL AMO F., SMITH D.E., SWIATEK P.J., GENDRON-MAGUIRE M.,
RA GREENSPAN R.J., MCMAHON A.P., GRIDLEY T.;
RT "Expression pattern of Notch, a mouse homolog of Drosophila Notch,
RT suggests an important role in early postimplantation mouse
RT development."
RL DEVELOPMENT 115:737-744(1992).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED ALMOST UNIFORMLY IN EARLY EMBRYOS.
CC -!- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC -!- SIMILARITY: CONTAINS 6 ANK REPEATS.
CC -!- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; Z11886; G288503;
CC MGD; MGI:97363; NOTCH1.
CC PROSITE; PS00010; ASX_HYDROXYL; 22.
CC PROSITE; PS00022; EGF_1; 34.
CC PROSITE; PS01186; EGF_2; 27.
CC PROSITE; PS01187; EGF_CA; 21.
CC PFAM; PF00008; EGF; 35.
CC PFAM; PF00023; ank; 6.
CC PFAM; PF00066; notch; 3.
CC HSP; P00740; LIXA.
CC DIFFERENTIATION; NEUROGENESIS; REPEAT; ANK REPEAT; EGF-LIKE DOMAIN;
CC TRANSMEMBRANE; SIGNAL; GLYCOPROTEIN.
FT SIGNAL 1 18
FT CHAIN 19 2531
FT DOMAIN 19 1725
FT TRANSMEM 1726 1746
FT DOMAIN 1747 2531
FT DOMAIN 24 1425
FT DOMAIN 1449 1462
FT DOMAIN 1445 1562
FT REPEAT 1445 1480
FT REPEAT 1481 1522
FT REPEAT 1523 1562
FT DOMAIN 1865 2075
FT REPEAT 1865 1910
FT REPEAT 1912 1942
FT REPEAT 1944 1975
FT ANK MOTIF 3.
-----
KW TRANSMEMBRANE; SIGNAL; GLYCOPROTEIN.
KW TRANSMEM 1726 1746
KW CHAIN 19 2531
KW SIGNAL 1 18
KW NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1.
KW EXTRACELLULAR (POTENTIAL).
KW POTENTIAL.
KW CYTOPLASMIC (POTENTIAL).
KW 36 X EGF-TYPE REPEATS.
KW CYS-RICH.
KW 3 X LIN/NOTCH REPEATS.
KW LIN/NOTCH 1.
KW LIN/NOTCH 2.
KW LIN/NOTCH 3.
KW 6 X ANK MOTIF REPEATS.
KW ANK MOTIF 1.
KW ANK MOTIF 2.
KW ANK MOTIF 3.
```

```
FT REPEAT 1978 2009
FT REPEAT 2011 2042
FT REPEAT 2044 2075
FT CARBOHYD 888 888
FT CARBOHYD 959 959
FT CARBOHYD 1179 1179
FT CARBOHYD 1241 1241
FT CARBOHYD 1489 1489
FT CARBOHYD 1587 1587
SQ SEQUENCE 2531 aa; 271312 MW; AD71189B CRC32;

Query Match 30.0%; Score 83; DB 1; Length 2531;
Best Local Similarity 41.7%; Pred. NO. 0.03;
Matches 15; Conservative 4; Mismatches 13; Indels 4; Gaps 1;

QY 9 DLAYCLNDCGCVVETLTGSHKHCKEGYQGVRC D 44
DB 1064 DSAPCKNGRCW---QTNTQYHCECRSGWTGVNCD 1095

RESULT 5
BTC_MOUSE STANDARD: PRT; 177 AA.
ID BTC_MOUSE
AC Q05928;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE BETACELLULIN PRECURSOR (BTC).
GN BTC OR BCN.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
[1]
SEQUENCE FROM N.A., AND SEQUENCE OF 32-54; 64-71 AND 75-111.
RN RP TISSUE=PANCREAS;
RX MEDLINE; 93206093.
RA SHING Y., CHRISTOFORI G., HANAHAN D., ONO Y., SASADA R.,
RA IGARASHI K., FOLKMAN J.;
RT "Betacellulin: a mitogen from pancreatic beta cell tumors."
RL SCIENCE 259:1604-1607(1993).
CC -!- FUNCTION: POTENT MITOGEN FOR RETINAL PIGMENT EPITHELIAL CELLS
CC AND VASCULAR SMOOTH MUSCLE CELLS. THE EFFECTS OF BETACELLULIN
CC ARE PROBABLY MEDIATED BY THE EGF RECEPTOR AND OTHER RELATED
CC RECEPTORS.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (PRECURSOR FORM);
CC EXTRACELLULAR (MATURE FORM).
CC -!- TISSUE SPECIFICITY: FOUND IN SEVERAL MOUSE TISSUES INCLUDING
CC KIDNEY, UTERUS AND LIVER AS WELL AS IN BETA TUMOR CELL LINE AND
CC MCF-7 CELLS. IT IS NOT DETECTED IN THE BRAIN.
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -!- SIMILARITY: STRONG, TO HUMAN BETACELLULIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; L08394; G293853;
CC PIR; A37408; A37408.
CC MGD; MGI:99439; BTC.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; 1.
CC PFAM; PF00008; EGF; 1.
CC HSP; P01135; 1YUT.
CC GROWTH FACTOR; MITOGEN; GLYCOPROTEIN; EGF-LIKE DOMAIN; TRANSMEMBRANE;
KW SIGNAL.
KW SIGNAL 1 31
FT CHAIN 32 111
FT PROPEP 112 177
FT BETACELLULIN.
FT REMOVED IN MATURE FORM.
```

FT DOMAIN 32 118 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 119 139 POTENTIAL.
 FT DOMAIN 140 177 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 146 153 ARG/LYS-RICH (BASIC).
 FT CARBOHYD 34 34 POTENTIAL.
 FT CARBOHYD 42 42 POTENTIAL.
 FT CARBOHYD 52 52 POTENTIAL.
 FT DOMAIN 65 105 EGF-LIKE.
 FT DISULFID 69 82 BY SIMILARITY.
 FT DISULFID 77 93 BY SIMILARITY.
 FT DISULFID 95 104 BY SIMILARITY.
 SQ SEQUENCE 177 AA; 19664 MW; B2A6D489 CRC32;

Query Match 29.8%; Score 82.5; DB 1; Length 177;
 Best Local Similarity 39.1%; Pred. No. 0.0031;
 Matches 18; Conservative 9; Mismatches 14; Indels 5; Gaps 3;

OY 1 HKPCRDKDLAYCNDGEC-FVIELTGTSHKCRKEGYGVRCDO 45
 DB 65 HFSRCQYKHYCIH-GRCFVDEQTPS---CICEKGYGARCER 106

RESULT 6
 ID GREFA_MYXVL STANDARD; PRT; 85 AA.
 AC P08072; 1988 (REL. 08, CREATED)
 DT 01-AUG-1988 (REL. 08, LAST SEQUENCE UPDATE)
 DT 01-AUG-1988 (REL. 08, LAST SEQUENCE UPDATE)
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
 DE GROWTH FACTOR (MGF).
 OS MYXOMA VIRUS (STRAIN LAUSANNE).
 OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXVIRIDAE; CHORDOPOXVIRINAE;
 OC LEPOPOXVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 87141350.
 RA UPTON C., MACEN J.L., MCFADDEN G.;
 RT "Mapping and sequencing of a gene from myxoma virus that is related
 RT to those encoding epidermal growth factor and transforming growth
 RT factor alpha.";
 RL J. VIROL. 61:1271-1275(1987).
 RN [2]
 RP SYNTHESIS, AND ACTIVITY.
 RX MEDLINE; 91182748.
 RA LIN Y.-Z., KE X.-H., TAM J.P.;
 RT "Synthesis and structure-activity study of myxoma virus growth
 RT factor.";
 RL BIOCHEMISTRY 30:3310-3314(1991).
 CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; M15806; G332300; --
 DR PIR; A26131; EGVM21.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; FALSE_NEG.
 DR PFAM; PF00008; EGF; 1.
 KW EGF-LIKE DOMAIN; GROWTH FACTOR; GLYCOPROTEIN.
 FT DOMAIN 33 77 EGF-LIKE.
 FT DISULFID 37 51
 FT DISULFID 45 65
 FT DISULFID 67 76
 FT CARBOHYD 48 48
 FT CARBOHYD 58 58
 SQ SEQUENCE 85 AA; 9629 MW; 195B4804 CRC32;

Query Match 29.6%; Score 82; DB 1; Length 85;
 Best Local Similarity 41.3%; Pred. No. 0.0019;
 Matches 19; Conservative 4; Mismatches 21; Indels 2; Gaps 2;

OY 3 KPCRDKDLAYCNDGECFVIELTGTSHKCRKEGYGVRCDOQL 47
 DB 35 KLCNDYKYNCLNGCTFTVALNNVSLNPFACACHINYVGSRC-QFI 79

RESULT 7
 ID SPIT_DROME STANDARD; PRT; 230 AA.
 AC Q01083;
 DT 01-NOV-1995 (REL. 32, CREATED)
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE PROTEIN SPITZ PRECURSOR.

GN SPI.
 OS DROSOPHILA MELANOGASTER (FRUIT FLY).
 OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;
 OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;
 OC DROSOPHILIDAE; DROSOPHILA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 92354912.
 RA RUTLEDGE B.J., ZHANG K., BIER E., JAN Y.N., PERRIMON N.;
 RT "The Drosophila spitz gene encodes a putative EGF-like growth factor
 RT involved in dorsal-ventral axis formation and neurogenesis.";
 RL GENES DEV. 6:1503-1517(1992).
 RN [2]
 RP FUNCTION.

RX MEDLINE; 95134691.
 RA FREEMAN M.;
 RT "The spitz gene is required for photoreceptor determination in the
 RT Drosophila eye where it interacts with the EGF receptor.";

CC -!- FUNCTION: PROBABLE LIGAND THAT IS BOUND BY THE RECEPTOR TOP.
 CC INVOLVED IN A NUMBER OF UNRELATED DEVELOPMENTAL CHOICES, FOR
 CC EXAMPLE, DORSAL-VENTRAL AXIS FORMATION, GLIAL MIGRATION, SENSORY
 CC ORGAN DETERMINATION, AND MUSCLE DEVELOPMENT. IT IS REQUIRED FOR
 CC PHOTORECEPTOR DETERMINATION.

CC -!- SUBCELLULAR LOCATION: TYPE I, MEMBRANE PROTEIN.
 CC -!- TISSUE SPECIFICITY: EXPRESSED THROUGHOUT THE EMBRYO.
 CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; M95199; G158455; --
 DR FLYBASE; FBgn0005672; spi.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PFAM; PF00008; EGF; 1.
 DR HSP; P01132; 1EPH.
 DR NEUROGENESIS; DEVELOPMENTAL PROTEIN; TRANSMEMBRANE; SIGNAL;
 KW GLYCOPROTEIN; EGF-LIKE DOMAIN.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 230 PROTEIN SPITZ.
 FT DOMAIN 18 139 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 140 160 POTENTIAL.
 FT DOMAIN 161 230 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 61 66 POLY-THR.
 FT DOMAIN 74 118 EGF-LIKE.
 FT DOMAIN 182 185 POLY-ASP.
 FT DISULFID 78 93 BY SIMILARITY.
 FT DISULFID 87 106 BY SIMILARITY.
 FT DISULFID 108 117 BY SIMILARITY.


```

FT DISULFID      77          93          BY SIMILARITY.
FT DISULFID      95         104          BY SIMILARITY.
SQ SEQUENCE     178 AA;   19746 MW;  7B8CE6B0 CRC32;

Query Match      29.1%; Score 80.5; DB 1; Length 178;
Best Local Similarity 41.3%; Pred. No. 0.0054;
Matches 19; Conservative 5; Mismatches 17; Indels 5; Gaps 3;

QY    1 HFKPCRDKLAYLNDGEC-FVIELTGLSHKHCRCKEGYGVRCDQ 45
      ||| :||: |||: ||| | ||| |||: |||:
Db    65 HFSRCPQYKHYCIK-GRCRFVVAEQTPS---CVCDEGYIGARCER 106

RESULT 9
GRFA_VACC      STANDARD; PRT; 142 AA.
ID AC P20494;
DT DT 01-FEB-1991 (REL. 17, CREATED)
DT DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE DE GROWTH FACTOR.
GN GN C1LR.
OS OS VACCINIA VIRUS (STRAIN COPENHAGEN).
OC OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXVIRIDAE; CHORDOPOXVIRINAE;
CC CC ORTHOPOXVIRUS.
RN RN [1]
RP RP SEQUENCE FROM N.A.
RX RX MEDLINE; 91021027.
RA RA GOEBEL S.-J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW J.P.,
RL RL PAOLETTI E.;
RL RL "The complete DNA sequence of vaccinia virus.";
RL RL VIROLOGY 179:247-266(1990).
RN RN [2]
RP RP COMPLETE GENOME.
RA RA GOEBEL S.-J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW J.P.,
RL RL PAOLETTI E.;
RL RL VIROLOGY 179:517-563(1990).
CC CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licensee@isb-sib.ch).
-----
CC CC EMBL; M35027; G335333; -.
DR DR PIR; C42503; WWVZ3C.
DR DR PROSITE; PS00022; EGF_1; 1.
DR DR PROSITE; PS01186; EGF_2; 1.
DR DR PFAM; PF00008; EGF; 1.
DR DR HSP; P01132; LEPH.
KW KW EGF-LIKE DOMAIN; GROWTH FACTOR; GLYCOPROTEIN.
FT FT DOMAIN 41 81
FT FT EGF-LIKE.
FT FT DISULFID 45 58
FT FT BY SIMILARITY.
FT FT DISULFID 53 69
FT FT BY SIMILARITY.
FT FT DISULFID 71 80
FT FT BY SIMILARITY.
FT FT CARBOHYD 34 34
FT FT POTENTIAL.
FT FT CARBOHYD 95 95
FT FT POTENTIAL.
SQ SEQUENCE 142 AA; 15777 MW; 00AFBC3 CRC32;

Query Match      28.7%; Score 79.5; DB 1; Length 142;
Best Local Similarity 38.5%; Pred. No. 0.0058;
Matches 15; Conservative 6; Mismatches 15; Indels 3; Gaps 2;

QY    5 CRKDKLAYLNDGECFVIELTGLSHKHCRCKEGYGVRVC 43
      | : |||: | : | : ||| | : |||
Db    45 CGPEGDGCYCLF-GDCIHARDIDG--MYCRGSHGYTGIRC 80

```


FT	DOMAIN	452	EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	537	552	BY SIMILARITY.
FT	DOMAIN	490	EGF-LIKE 13, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	554	563	BY SIMILARITY.
FT	DOMAIN	526	EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	570	580	BY SIMILARITY.
FT	DOMAIN	564	EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	575	589	BY SIMILARITY.
FT	DOMAIN	566	EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	591	600	BY SIMILARITY.
FT	DOMAIN	603	EGF-LIKE 17, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	607	618	BY SIMILARITY.
FT	DOMAIN	641	EGF-LIKE 18, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	612	627	BY SIMILARITY.
FT	DOMAIN	678	EGF-LIKE 19, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	629	638	BY SIMILARITY.
FT	DOMAIN	716	EGF-LIKE 20, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	645	655	BY SIMILARITY.
FT	DOMAIN	753	EGF-LIKE 21, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	650	664	BY SIMILARITY.
FT	DOMAIN	791	EGF-LIKE 22, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	666	675	BY SIMILARITY.
FT	DOMAIN	827	EGF-LIKE 23, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	682	693	BY SIMILARITY.
FT	DOMAIN	869	EGF-LIKE 24, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	687	702	BY SIMILARITY.
FT	DOMAIN	899	EGF-LIKE 25, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	704	713	BY SIMILARITY.
FT	DOMAIN	945	EGF-LIKE 26, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	720	730	BY SIMILARITY.
FT	DOMAIN	983	EGF-LIKE 27, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	725	739	BY SIMILARITY.
FT	DOMAIN	1021	EGF-LIKE 28, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	741	750	BY SIMILARITY.
FT	DOMAIN	1059	EGF-LIKE 29, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	757	768	BY SIMILARITY.
FT	DOMAIN	1097	EGF-LIKE 30, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	762	777	BY SIMILARITY.
FT	DOMAIN	1145	EGF-LIKE 31, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	779	788	BY SIMILARITY.
FT	DOMAIN	1183	EGF-LIKE 32, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	795	806	BY SIMILARITY.
FT	DOMAIN	1221	EGF-LIKE 33, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	800	815	BY SIMILARITY.
FT	DOMAIN	1265	EGF-LIKE 34, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	817	826	BY SIMILARITY.
FT	DOMAIN	1307	EGF-LIKE 35, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	833	844	BY SIMILARITY.
FT	DOMAIN	1346	EGF-LIKE 36, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	838	855	BY SIMILARITY.
FT	DOMAIN	1384	CYS-RICH.	FT	DISULFID	857	866	BY SIMILARITY.
FT	DOMAIN	1387	6 X ANK MOTIF REPEATS.	FT	DISULFID	873	884	BY SIMILARITY.
FT	DOMAIN	1426	ANK MOTIF 1.	FT	DISULFID	878	893	BY SIMILARITY.
FT	DOMAIN	1462	ANK MOTIF 2.	FT	DISULFID	895	904	BY SIMILARITY.
FT	DOMAIN	1865	ANK MOTIF 3.	FT	DISULFID	911	922	BY SIMILARITY.
FT	DOMAIN	1865	ANK MOTIF 4.	FT	DISULFID	916	931	BY SIMILARITY.
FT	DOMAIN	1944	ANK MOTIF 5.	FT	DISULFID	933	942	BY SIMILARITY.
FT	DOMAIN	1978	ANK MOTIF 6.	FT	DISULFID	987	998	BY SIMILARITY.
FT	DOMAIN	2011	BY SIMILARITY.	FT	DISULFID	992	1007	BY SIMILARITY.
FT	DOMAIN	2042	BY SIMILARITY.	FT	DISULFID	1009	1018	BY SIMILARITY.
FT	DOMAIN	2044	BY SIMILARITY.	FT	DISULFID	1025	1036	BY SIMILARITY.
FT	DOMAIN	24	BY SIMILARITY.	FT	DISULFID	1030	1045	BY SIMILARITY.
FT	DOMAIN	31	BY SIMILARITY.	FT	DISULFID	1047	1056	BY SIMILARITY.
FT	DOMAIN	46	BY SIMILARITY.	FT	DISULFID	1063	1074	BY SIMILARITY.
FT	DOMAIN	57	BY SIMILARITY.	FT	DISULFID	1068	1083	BY SIMILARITY.
FT	DOMAIN	63	BY SIMILARITY.	FT	DISULFID	1085	1094	BY SIMILARITY.
FT	DOMAIN	68	BY SIMILARITY.	FT	DISULFID	1101	1122	BY SIMILARITY.
FT	DOMAIN	87	BY SIMILARITY.	FT	DISULFID	1116	1131	BY SIMILARITY.
FT	DOMAIN	89	BY SIMILARITY.	FT	DISULFID	1133	1142	BY SIMILARITY.
FT	DOMAIN	106	BY SIMILARITY.	FT	DISULFID	1149	1160	BY SIMILARITY.
FT	DOMAIN	111	BY SIMILARITY.	FT	DISULFID	1154	1169	BY SIMILARITY.
FT	DOMAIN	129	BY SIMILARITY.	FT	DISULFID	1171	1180	BY SIMILARITY.
FT	DOMAIN	144	BY SIMILARITY.	FT	DISULFID	1187	1198	BY SIMILARITY.
FT	DOMAIN	155	BY SIMILARITY.	FT	DISULFID	1192	1207	BY SIMILARITY.
FT	DOMAIN	164	BY SIMILARITY.	FT	DISULFID	1209	1218	BY SIMILARITY.
FT	DOMAIN	166	BY SIMILARITY.	FT	DISULFID	1225	1244	BY SIMILARITY.
FT	DOMAIN	182	BY SIMILARITY.	FT	DISULFID	1238	1253	BY SIMILARITY.
FT	DOMAIN	189	BY SIMILARITY.	FT	DISULFID	1255	1264	BY SIMILARITY.
FT	DOMAIN	206	BY SIMILARITY.	FT	DISULFID	1271	1284	BY SIMILARITY.
FT	DOMAIN	222	BY SIMILARITY.	FT	DISULFID	1276	1293	BY SIMILARITY.
FT	DOMAIN	227	BY SIMILARITY.	FT	DISULFID	1295	1304	BY SIMILARITY.
FT	DOMAIN	243	BY SIMILARITY.	FT	DISULFID	1311	1322	BY SIMILARITY.
FT	DOMAIN	254	BY SIMILARITY.	FT	DISULFID	1316	1334	BY SIMILARITY.
FT	DOMAIN	272	BY SIMILARITY.	FT	DISULFID	1336	1345	BY SIMILARITY.
FT	DOMAIN	281	BY SIMILARITY.	FT	DISULFID	1352	1363	BY SIMILARITY.
FT	DOMAIN	283	BY SIMILARITY.	FT	DISULFID	1357	1372	BY SIMILARITY.
FT	DOMAIN	299	BY SIMILARITY.	FT	DISULFID	1374	1383	BY SIMILARITY.
FT	DOMAIN	306	BY SIMILARITY.	FT	DISULFID	1391	1403	BY SIMILARITY.
FT	DOMAIN	323	BY SIMILARITY.					
FT	DOMAIN	339	BY SIMILARITY.					
FT	DOMAIN	350	BY SIMILARITY.					
FT	DOMAIN	359	BY SIMILARITY.					
FT	DOMAIN	370	BY SIMILARITY.					
FT	DOMAIN	376	BY SIMILARITY.					
FT	DOMAIN	381	BY SIMILARITY.					
FT	DOMAIN	398	BY SIMILARITY.					
FT	DOMAIN	400	BY SIMILARITY.					
FT	DOMAIN	416	BY SIMILARITY.					
FT	DOMAIN	423	BY SIMILARITY.					
FT	DOMAIN	440	BY SIMILARITY.					
FT	DOMAIN	456	BY SIMILARITY.					
FT	DOMAIN	467	BY SIMILARITY.					
FT	DOMAIN	478	BY SIMILARITY.					
FT	DOMAIN	494	BY SIMILARITY.					
FT	DOMAIN	499	BY SIMILARITY.					
FT	DOMAIN	516	BY SIMILARITY.					
FT	DOMAIN	525	BY SIMILARITY.					
FT	DOMAIN	532	BY SIMILARITY.					

Query Match 28.5%; Score 79; DB 1; Length 2531;

Best Local Similarity 38.9%; Pred. No. 0.089;

Matches 14; Conservative 5; Mismatches 13; Indels 4; Gaps 1;

QY 9 DLAYCLNDCGEFVIELTGTSHKHCRCKEGYQGVRC 44

DB 1064 DSAPCKNGKCN---QTNTQYHCECRSGWTGFNC 1095

CC		EMBL; X04571; G31121; -.	
DR	PIR;	A25531; ECHU.	
DR	MIM;	131530; -.	
DR	MIM;	191810; -.	
DR	PROSITE;	PS00010; ASX_HYDROXYL; 3.	
DR	PROSITE;	PS00022; EGF_1; 1.	
DR	PROSITE;	PS01186; EGF_2; 7.	
DR	PROSITE;	PS01187; EGF_CA; 3.	
DR	PFAM;	PF00008; EGF; 8.	
DR	PFAM;	PF00058; ldl_recept_b; 6.	
DR	HSSP;	PO1132; LEPH.	
DR	EGF-LIKE DOMAIN;	REPEAT; GROWTH FACTOR; TRANSMEMBRANE; GLYCOPROTEIN;	
KW	SIGNAL;	POLYMORPHISM.	
FT	SIGNAL	1	POTENTIAL.
FT	CHAIN	23	EPIDERMAL GROWTH FACTOR.
FT	DOMAIN	23	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	1033	POTENTIAL.
FT	DOMAIN	1054	CYTOPLASMIC (POTENTIAL).

INTERFERENCE

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 1999, 00:32:25 ; Search time 10.18 Seconds
(without alignments)
52.765 Million cell updates/sec

Title: US-09-107-979-4
Perfect score: 277
Sequence: 1 HFKPCRDKLAYCLNDGEF.....SHKCRCKEGYQGVRCDOFL 47

Scoring table: BLOSUM62

Searched: 119832 seqs, 11428610 residues

Database : Issued Patents AA: *
1: /cg2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cg2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cg2_6/ptodata/1/iaa/PCTUS9_COMB.pep.*
4: /cg2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	116.5	42.1	52	1	US-08-417-640A-1
2	116.5	42.1	52	1	US-08-760-815-1
3	116.5	42.1	52	2	US-08-761-038-1
4	113.5	41.0	669	1	US-07-847-743B-8
5	113.5	41.0	66	1	US-07-847-743B-10
6	113.5	41.0	669	1	US-07-847-743B-13
7	113.5	41.0	95	1	US-07-847-743B-14
8	113.5	41.0	625	1	US-07-847-743B-26
9	113.5	41.0	54	1	US-08-179-481-111
10	113.5	41.0	669	1	US-08-456-201-8
11	113.5	41.0	66	1	US-08-456-201-10
12	113.5	41.0	669	1	US-08-456-201-13
13	113.5	41.0	95	1	US-08-456-201-14
14	113.5	41.0	625	1	US-08-456-201-26
15	113.5	41.0	52	1	US-08-417-640A-3
16	113.5	41.0	52	1	US-08-760-815-3
17	113.5	41.0	669	2	US-08-330-161-11
18	113.5	41.0	95	2	US-08-330-161-12
19	113.5	41.0	669	2	US-08-456-241-8
20	113.5	41.0	66	2	US-08-456-241-10
21	113.5	41.0	669	2	US-08-456-241-13
22	113.5	41.0	95	2	US-08-456-241-14
23	113.5	41.0	625	2	US-08-456-241-26
24	113.5	41.0	669	2	US-08-440-401-11
25	113.5	41.0	95	2	US-08-440-401-12
26	113.5	41.0	669	2	US-08-419-878B-11
27	113.5	41.0	95	2	US-08-419-878B-12
28	113.5	41.0	52	2	US-08-761-038-3
29	113.5	41.0	669	3	PCT-US92-04295A-8
30	113.5	41.0	66	3	PCT-US92-04295A-10
31	113.5	41.0	669	3	PCT-US92-04295A-13
32	113.5	41.0	95	3	PCT-US92-04295A-14
33	113.5	41.0	625	3	PCT-US92-04295A-26
34	111.5	40.3	54	1	US-08-179-481-110
35	108.5	39.2	50	1	US-08-036-555B-152
36	108.5	39.2	50	1	US-08-459-569-152
37	108.5	39.2	45	1	US-08-168-091A-39
38	108.5	39.2	50	1	US-08-249-322A-152
39	108.5	39.2	50	2	US-08-469-526A-152

ALIGNMENTS

```
RESULT 1
US-08-417-640A-1
; Sequence 1, Application US/08417640A
; Patent No. 5670342
; GENERAL INFORMATION:
; APPLICANT: Carnahan, Josette F.
; APPLICANT: Hara, Shinichi
; APPLICANT: Lu, Hsieng S.
; APPLICANT: Mayer, John P.
; APPLICANT: Yoshinaga, Steven K.
; TITLE OF INVENTION: NDF Peptides
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/417,640A
; FILING DATE:
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Mazza, Richard J.
; REFERENCE/DOCKET NUMBER: A-310
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 52 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-417-640A-1

Query Match 42.1%; Score 116.5; DB 1; Length 52;
Best Local Similarity 33.3%; Pred. No. 9.8e-08;
Matches 16; Conservative 15; Mismatches 16; Indels 1; Gaps 1;

QY 1 HFKPCRDKLAYCLNDGEFVIETLTGSHKH-CRCKEGYQGVRCDOFL 47
Db 2 HLYKCAEKETFCVNGCEFWKDLNPNRYLCKQCPGFTGRCQNYV 49.

RESULT 2
US-08-760-815-1
; Sequence 1, Application US/08760815
; Patent No. 5686413
; GENERAL INFORMATION:
; APPLICANT: Carnahan, Josette F.
; APPLICANT: Hara, Shinichi
; APPLICANT: Lu, Hsieng S.
; APPLICANT: Mayer, John P.
; APPLICANT: Yoshinaga, Steven K.
; TITLE OF INVENTION: NDF Peptides
; NUMBER OF SEQUENCES: 6
```

```

; NAME: Mazza, Richard J.
; REFERENCE/DOCKET NUMBER: A-310
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 52 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-761-038-1

Query Match 42.1%; Score 116.5; DB 2; Length 52;
Best Local Similarity 33.3%; Pred. No. 9.8e-08;
Matches 16; Conservative 15; Mismatches 16; Indels 1

QY 1 HFKPRCKDLAVCLNDGECFVETLTGSHKH-CRCKEYQGVRCDFL 47
| | : : : : | | | | : : : : | : | : | : | :
Db 2 HLVKCAEKETFCVNGEGCFMWKDLNFSRYLCKQPGFTGARCONYV 49

RESULT 4
US-07-847-743B-8
; Sequence 8, Application US/07847743B
; Patent No. 5367060
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: Structure, Production and Use of
; TITLE OF INVENTION: Heregulin
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/847,743B
; FILING DATE: 19920306
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/705256
; FILING DATE: 24-May-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/765212
; FILING DATE: 25-Sep-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/790801
; FILING DATE: 08-No. 5367060-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hensley, Max D.
; REGISTRATION NUMBER: 27,043
; REFERENCE/DOCKET NUMBER: 712p3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/266-1489
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 669 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; US-07-847-743B-8

Query Match 41.0%; Score 113.5; DB 1; Length 669;
Best Local Similarity 34.8%; Pred. No. 3.1e-06;

```


Query Match 41.0%; Score 113.5; DB 1; Length 54;
Best Local Similarity 34.8%; Pred. NO. 2.3e-07;


```
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456,201
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/126,145
FILING DATE:
APPLICATION NUMBER: 07/880,917
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/705256
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/765212
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/790801
ATTORNEY/AGENT INFORMATION:
NAME: Hensley, Max D.
REGISTRATION NUMBER: 27,034
REFERENCE/DOCKET NUMBER: 712P4
TELEPHONE: 415/266-1489
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 669 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-456-201-13

Query Match 41.0%; Score 113.5; DB 1; Length 669;
Best Local Similarity 34.8%; Pred. No. 3.1e-06;
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

QY 1 HFKPCDKDLAYCLNDGECFVETLTGSHKH-CRCKEGYQGVRCDO 45
| | | | | : | | | | | : | | | | | : | | | | | :
Db 222 HLVKCAEKEKTCVNGGECFMVKDLSNPRLCKQPGTGARCTE 267

RESULT 13
US-08-456-201-14
Sequence 14, Application US/08456201
Patent No. 5641869
GENERAL INFORMATION:
APPLICANT: Vandlen, Richard L.,
APPLICANT: Holmes, William E.,
TITLE OF INVENTION: Structure, Production and Use of
OPERATING SYSTEM: Herregulin 2 Ligands
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456,201
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/126,145
FILING DATE:
APPLICATION NUMBER: 07/880,917
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/705256
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/765212
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/790801
ATTORNEY/AGENT INFORMATION:
NAME: Hensley, Max D.
REGISTRATION NUMBER: 27,034
REFERENCE/DOCKET NUMBER: 712P4
FILING DATE:
TELECOMMUNICATION INFORMATION:

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456,201
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/126,145
FILING DATE:
APPLICATION NUMBER: 07/880,917
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/705256
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/765212
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/790801
ATTORNEY/AGENT INFORMATION:
NAME: Hensley, Max D.
REGISTRATION NUMBER: 27,034
REFERENCE/DOCKET NUMBER: 712P4
FILING DATE:
TELECOMMUNICATION INFORMATION:
```

Search completed: October 14, 1999, 00:43:10
Job time: 645 sec

THIS PAGE BLANK (USPTO)